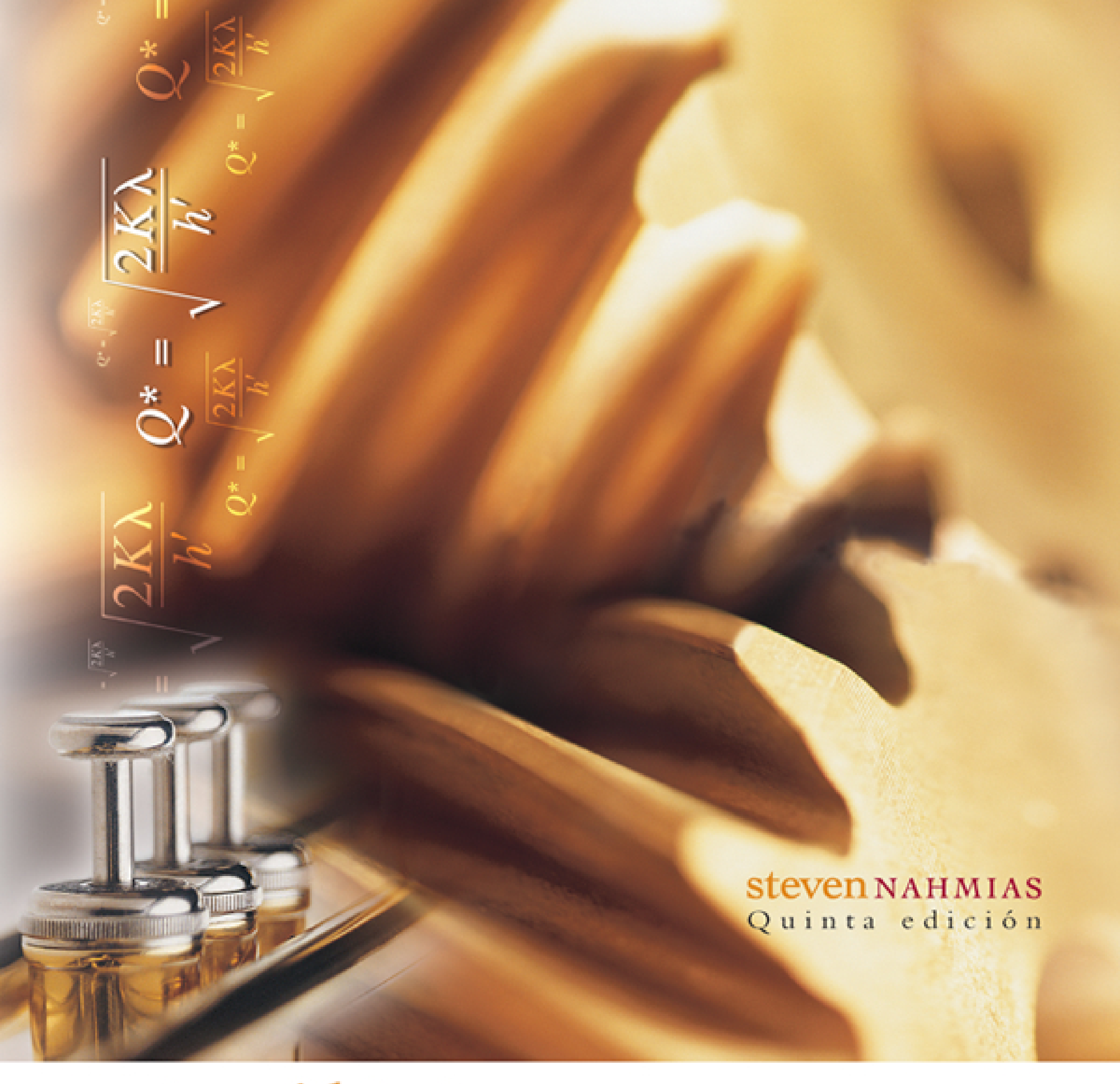


Análisis de la producción y las operaciones

McGraw Hill

INCLUYE CD-ROM

La matriz producto-proceso

Estructura del proceso Etapa del ciclo de vida del proceso	I Bajo volumen, baja estandarización, único	II Productos múltiples, bajo volumen	III Algunos productos principales, mayor volumen	IV Alto volumen, alta estandarización, productos masivos
I Flujo mezclado (taller por tareas)	Imprenta comercial			Vacío
II Flujo de línea desconectada (lote)		Maquinaria pesada		
III Flujo de línea conectada (línea de ensamble)			Ensambladora de autos	
IV Flujo continuo	Vacío			Refinería de azúcar

Fuente: Robert H. Hayes y Steven C. Wheelwright, "Link Manufacturing Process and Product Life Cycles", en *Harvard Business Review*, enero-febrero de 1979.

Análisis de la producción y las operaciones

Análisis de la producción y las operaciones

Quinta edición

Steven Nahmias
Santa Clara University

Revisores técnicos
Alejandro Terán Castellanos
Instituto Tecnológico Autónomo de México

Elmer J. Murrieta Murrieta
Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Cuernavaca

Filiberto González Hernández
Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Estado de México



MÉXICO • AUCKLAND • BOGOTÁ • BUENOS AIRES • CARACAS • GUATEMALA
LISBOA • LONDRES • MADRID • MILÁN • MONTREAL • NUEVA DELHI • NUEVA YORK
SAN FRANCISCO • SAN JUAN • SAN LUIS • SANTIAGO
SÃO PAULO • SIDNEY • SINGAPUR • TORONTO

Director Higher Education: Miguel Ángel Toledo Castellanos
Director editorial: Ricardo del Bosque Alayón
Editor sponsor: Pablo Eduardo Roig Vázquez
Editora de desarrollo: Ana Laura Delgado Rodríguez
Supervisor de producción: Zeferino García García

Traducción: Raúl Arrioja Juárez
Birgit Nuding Fleischmann
Jorge Yescas Milanés

ANÁLISIS DE LA PRODUCCIÓN Y LAS OPERACIONES



Edificio Punta Santa Fe
Prolongación Paseo de la Reforma 1015 Torre A, Piso 17,
Colonia Desarrollo Santa Fe, Delegación Álvaro Obregón
C.P. 01376, México, D.F.
Miembro de la Cámara Nacional de la Industria Editorial Mexicana, Reg. Núm. 736

ISBN-13: 978-970-10-6239-5
ISBN-10: 970-10-6239-6

Traducido de la quinta edición de: *Production and Operations Analysis.*

ISBN: 007-286538-5

1234567890 09865432107

Impreso en México *Printed in Mexico*

A la memoria de mis padres
Liza y Morris

Steven Nahmias

Sumario

Contenido

Capítulo 3
Planeación agregada 108

Suplemento 1 Programación lineal 154

Capítulo 8
Programación de operaciones 401

Suplemento 2 Teoría de colas 457

Capítulo 9
Programación de proyectos 484

Capítulo 10
Distribución y ubicación de instalaciones 535

Acerca del autor

Steven Nahmias es profesor de Sistemas de Información para las Operaciones y la Administración en la Leavey School of Bussiness and Administration de la Universidad de Santa Clara. Tiene una licenciatura en Matemáticas y Física del Queens College, y una licenciatura en Ingeniería Industrial de la Universidad de Columbia; cursó la maestría y el doctorado en Investigación de las Operaciones en la Universidad Northwestern. Asimismo, ha sido profesor de la Universidad de Pittsburgh, del Instituto Tecnológico de Georgia y de la Universidad de Stanford.

El doctor Nahmias es ampliamente conocido por sus investigaciones sobre modelos estocásticos para inventarios, y es autor y coautor de más de 50 artículos científicos publicados en diversas revistas especializadas. Ha sido editor de *Operations Research*, en el área de Logística de la Cadena de Abastecimiento; editor asociado de *Management Science*; editor asociado de *Naval Research Logistics*, y actualmente se desempeña como editor en jefe de *M&SOM*. Recibió el primer lugar en la 39 Competencia Nicholson para el Mejor Ensayo Estudiantil (1971), segundo lugar en la Competencia TIMS para el Mejor Ensayo Estudiantil (1972), y obtuvo el Premio Universitario a la Excelencia Académica que otorga la Universidad de Santa Clara (1998).

Además de sus actividades académicas, el doctor Nahmias ha fungido como consultor de diferentes compañías y dependencias, como Litton Industries, Xerox Corporation, Lex Automotive y el Departamento de Transporte del Condado de Santa Clara. Entre sus diferentes pasatiempos se encuentran el golf y el ciclismo, y toca la trompeta con varias bandas de jazz de su comunidad.

Prefacio a la quinta edición

Con el objetivo de beneficiar a instructores y a estudiantes, esta edición incluye numerosas modificaciones pedagógicas, aunque se ha conservado la misma estructura, títulos y secuencia, de los capítulos; asimismo, los casi 550 ejercicios permanecen esencialmente igual que en la cuarta edición.

- En esta edición se incluyen resúmenes al inicio de cada capítulo, con el fin de proporcionar al estudiante un panorama general del tema antes de que se profundice en los pormenores. Mi recomendación es que todos los lectores dediquen tiempo a los resúmenes que se encuentran en la parte inicial, ya que los resúmenes proporcionan un panorama general del flujo lógico del material tratado en los capítulos.
- Además, hay otros rasgos que caracterizan la presente edición. En muchos de los capítulos aparecen nuevas Aplicaciones instantáneas. Asimismo, las introducciones a todos los capítulos han sido reescritas con "ganchos" que permiten captar la atención de los lectores. (Agradezco a Dave Luenberger por haberme señalado el valor de tales anzuelos.) Muchos de los temas y cuestiones susceptibles de cambio han sido actualizados a lo largo de todo el libro.

Agradezco a todos aquellos estudiantes y colegas que me mandaron mensajes de correo electrónico con sugerencias sobre modificaciones y complementos que les gustaría ver. Desafortunadamente, fue imposible incluir todas sus recomendaciones y temas de interés. Sin embargo, al revisar el libro consideré cuidadosamente todos y cada uno de los comentarios que me fueron enviados. Los exhorto a que sigan comunicándose conmigo para hacerme saber los aspectos del libro que pueden mejorarse, así como los temas que debieran incluirse. Mi dirección de correo electrónico es snahmias@scu.edu. Muchas gracias de nuevo por su continuo apoyo.

Por último, me gustaría agradecer a las siguientes personas por su apoyo: Joseph Geunes, de la Universidad de Florida; Dennis Kroll, de la Universidad de Bradley; Alan Leduc, de la Universidad de Ball State; Dennis Webster, de la Universidad Estatal de Louisiana, y Brian Craig, de la Universidad de Lamar.

Introducción

1 LA FUNCIÓN DE LA ADMINISTRACIÓN DE LA PRODUCCIÓN Y LAS OPERACIONES

Este libro trata sobre la producción en un sentido amplio: esto es, la producción tanto de *artículos* como de *servicios*. El término producción tiene sus raíces en los estudios de automatización de fábricas; el término *administración de las operaciones* se utiliza para reflejar aspectos más amplios. Este libro trata sobre los métodos analíticos que se utilizan para apoyar la administración de la producción y las operaciones, es decir, el proceso de dirigir personas y recursos para crear un producto o servicio. Las operaciones se relacionan con la logística que apoya la función de la producción.

Las finanzas, la mercadotecnia y la producción son las tres áreas funcionales de una compañía. La mercadotecnia actúa como un intermediario con el mundo exterior, elige la forma en la que se presentan los productos al consumidor y explora hacia dónde se dirigen las preferencias de éstos. Las finanzas son responsables de encontrar fuentes de financiamiento externo y administran el capital dentro de la compañía. La producción se encarga de convertir las materias primas en productos o de proporcionar un conjunto específico de servicios. Para que una compañía opere con eficiencia, las actividades de estas tres áreas funcionales deben coordinarse muy cuidadosamente.

Bowman y Fetter (1967) enumeran los siguientes problemas económicos como los más relevantes para la administración de la producción:

1. Inventarios.
2. Control y programación de la producción.
3. Selección y reemplazo de equipo.
4. Mantenimiento.
5. Tamaño y ubicación de plantas.
6. Estructura y disposición de la planta.
7. Inspección y control de calidad.
8. Manejo del tráfico y de los materiales.
9. Métodos.

Aunque esta lista se recopiló hace 30 años, aún representa con precisión las decisiones económicas más importantes de la administración de la producción. En este libro se revisarán todas estas áreas, a excepción de las dos últimas. Aunque el manejo de materiales y los métodos de trabajo son temas relevantes, no son áreas en las que los métodos analíticos hayan tenido un impacto significativo.

2 HISTORIA DE LA ADMINISTRACIÓN DE LA PRODUCCIÓN

De acuerdo con Skinner (1985), hay cinco periodos de la historia industrial que sobresalen en el desarrollo de la administración de la manufactura:

1780-1850	Los líderes manufactureros como capitalistas de la tecnología.
1850-1890	Los líderes manufactureros como arquitectos de la producción en masa.
1890-1920	La administración de la manufactura permea a la organización.

1920-1960 La administración de la manufactura refina sus habilidades de control y estabilidad.
1960-1980 Sacudida de los aspectos fundamentales de la administración industrial.

Durante los primeros años de la Revolución Industrial, la producción comenzó a cambiar de actividades de bajo volumen a operaciones de mayor escala. Aunque la escala de las operaciones iniciales tenía un tamaño importante, la maquinaria no era particularmente compleja y las operaciones de producción eran rígidas. La administración de estas operaciones permaneció fundamentalmente en manos de la alta dirección, con la ayuda de supervisores. Con frecuencia, las condiciones de trabajo durante este periodo eran muy malas.

El primer empuje de la Revolución Industrial se llevó a cabo entre 1850-1890. En esa época surgieron los conceptos *producción en masa* y *línea de ensamble*. Como el carbón ya podía transportarse con mucha eficiencia, era posible situar plantas en diversas ubicaciones. En ese entonces, el capataz de la planta tenía un enorme poder e influencia.

Skinner afirma que la labor del gerente de producción surgió entre 1890 y 1920. Los procesos de manufactura se volvieron demasiado complejos como para que los siguiera manejando la alta dirección. Esta complejidad ocasionó que fueran necesarias técnicas científicas de administración. A Frederick Taylor (también llamado el padre de la ingeniería industrial) se le atribuye el concepto *administración científica*. La mayoría de las técnicas de la administración científica de principios de siglo implicaban únicamente la fragmentación de una tarea en diversos componentes. Estas técnicas eran quizá más de ordenamiento que científicas. Con los nuevos niveles de complejidad, el capataz de una planta ya no pudo coordinar la producción de una línea de producción variada o de ajustar los calendarios de producción. Entonces llegó la era del especialista.

Pese a la enorme depresión mundial que se registró en los años treinta, el periodo comprendido entre 1920 y 1960 puede considerarse una era dorada para el desarrollo de la industria en Estados Unidos. Para 1960, Estados Unidos se había convertido en la mayor potencia económica mundial. Al crecer el movimiento laboral, las condiciones de trabajo mejoraron enormemente y comenzaron a llegar a las fábricas auténticos métodos científicos. Los modelos matemáticos para el aprendizaje, el control de inventarios, el control de calidad, la programación de la producción y la administración de proyectos ganaron aceptación entre la comunidad de usuarios. Durante ese periodo, la alta dirección se convierte en un objetivo de los profesionales de la producción.

A partir de 1960, muchas compañías estadounidenses perdieron el dominio que tenían en ciertos mercados. Los productos que tradicionalmente se generaban en Estados Unidos, ahora se importan de Alemania, de Japón o del Lejano Oriente: a veces se produce de manera más económica y con mayor calidad; además de que las relaciones entre el empleado y la administración son por lo general mejores. Por ejemplo, los círculos de calidad, introducidos en Japón, permitieron a los empleados opinar sobre el desarrollo del producto y los procedimientos de producción. En Japón, más que en otros países, se adoptaron métodos de producción científica muy sofisticados. En ese país existe un gran número de robots y sistemas de manufactura flexibles y modernos.

Ahora los gerentes de producción requieren una mayor capacitación y preparación. Tienen que estar mejor versados en diversas tecnologías y mostrarse capaces de manejar sistemas de información y soporte complejos. Las computadoras continuarán mejorando y conforme la capacidad del cómputo llegue a las fábricas, los modelos matemáticos jugarán un papel cada vez más determinante en la administración actual de la producción y las operaciones. El objetivo de este texto es proporcionar al potencial profesional de la producción todas las herramientas técnicas necesarias para enfrentar el futuro inmediato.

3 PANORAMA GENERAL DEL TEXTO

Albert Einstein dijo que la física debe explicarse de una forma sencilla, pero nunca simple. Esto refleja la filosofía de este libro.

Las características de este libro incluyen

- Extensión.
- Profundidad.
- Claridad.
- Modularidad.
- Ejemplos.
- Reforzamiento.
- Notas históricas y resúmenes.
- Problemas en hojas de cálculo.
- Aplicaciones instantáneas.

La mayor parte de los capítulos incluye material que raramente se encuentra en textos de investigación, además de tópicos que sólo se han tratado en artículos de revistas especializadas. La escritura es concisa, pero todos los conceptos se explican cuidadosamente. Los capítulos están diseñados para poderse manejar de manera independiente. Algunas secciones pueden alternarse sin perder la continuidad, lo que permite al instructor ajustar la secuencia y profundidad de cobertura a sus necesidades. He hecho un amplísimo uso de ejemplos; algunos de ellos son ejemplos de caso que se cubren en más de un capítulo, mientras que otros son ejemplos breves utilizados para ilustrar una técnica o concepto en particular. He tratado de presentar ejemplos realistas, que muestren al estudiante la forma en que un método puede usarse en la práctica.

He incluido problemas al final de cada sección, así como al final de la mayor parte de los capítulos. Los problemas de sección tienen dos ventajas: la primera es que proporcionan al estudiante un refuerzo inmediato del material comentado en esa sección en particular; la segunda es que facilitará al profesor a cubrir parte de los capítulos y asignar problemas adecuados. Hay más de 550 problemas en el libro.

Cada capítulo, excepto el primero, incluye una sección denominada "Notas históricas". Como estudiante, siempre me encontré con libros de texto en cierta forma austeros, los cuales transmitían muy poco de la historia de los temas que en ellos se desarrollaban. Pienso que es interesante saber cuándo y cómo ocurrieron las cosas. Me doy cuenta de que afronto un gran riesgo al incluir estas secciones, ya que puedo haber identificado incorrectamente algunas de las fuentes originales. Agradecería a los lectores conocedores que así me lo hicieran saber, si tal fuere el caso. Los resúmenes que se encuentran al final de cada capítulo proporcionarán al lector un panorama general del material que se trata en los capítulos.

Se incluyeron problemas para ser diseñados con la ayuda de una computadora (en particular con un programa para hojas de cálculo). Estos problemas están marcados con una gráfica para identificarlos y casi siempre aparecen en el apartado de problemas que se encuentra al final de los capítulos.

La forma en que debe distribuirse la secuencia de los diferentes temas es una cuestión personal. Mi fundamento es el siguiente: el primer capítulo es principalmente expositivo y proporciona una introducción no técnica para las cuestiones estratégicas importantes con las que se enfrentan hoy en día todas las compañías. El resto del libro cubre tres áreas generales: inventarios (definidos en un sentido amplio), programación y diseño de instalaciones, y calidad. Las secciones de inventarios abarcan los capítulos 2-6, en los que se cubren pronósticos, planeación agregada, sistemas determinísticos y estocásticos para un solo artículo, y cadenas de abastecimiento, respectivamente. MRP y JIT, vistos en el capítulo 7, se tratan tanto con inventarios como con programación, y proporcionan el enlace para los siguientes tres capítulos. Éstos son los capítulos 8-10, que incluyen la programación de taller para tareas, programación de proyectos, y disposición y ubicación. Los dos capítulos finales tratan sobre calidad, definida en un sentido amplio.

El capítulo 11 cubre de forma extensa el control estadístico de calidad, el muestreo de aceptación, y varios tópicos adicionales de calidad. El capítulo 12 difícilmente se encuentra en los textos de operaciones. Pienso que un libro de texto sobre calidad debe incluir también confiabilidad y mantenimiento.

Las ediciones previas del libro han tenido gran éxito y se utilizan tanto en escuelas de ingeniería como de negocios. Aunque el libro no ha perdido su carácter cuantitativo, las revisiones han ayudado a ampliar la audiencia y a sumar material cualitativo. El nuevo capítulo de la cadena de abastecimiento incluye varias secciones cuantitativas y está enfocado básicamente a cuestiones estratégicas. Debido al diseño modular del libro, puede adaptarse fácilmente a las necesidades del instructor. Los primeros seis capítulos pueden servir como un curso razonable para un trimestre sobre inventarios y temas relacionados; los capítulos 7-10 proporcionan un tratamiento razonable sobre programación, incluyendo MRP y JIT; los dos últimos capítulos pueden formar la base para un curso de calidad y confiabilidad. Por otro lado, puede conformarse un curso panorámico, tomando diversas secciones de cada capítulo. Creo que existe suficiente material en el texto para una secuencia de dos semestres, o dos trimestres, también.

4 PANORAMA GENERAL DE CADA CAPÍTULO

Capítulo 1. Estrategia y competencia

El capítulo 1 proporciona un breve panorama general de las cuestiones relevantes para la estrategia de operaciones y la competencia. La introducción posiciona a la estrategia de manufactura como parte de la estrategia general de la compañía. La sección 1 considera la cuestión de nuestra evolución hacia una economía de servicio, y las preocupaciones que esto puede traer en el futuro. La sección 2 proporciona un panorama general para el marco de trabajo en el que deben abordarse cuestiones estratégicas. La sección 3 considera la clásica visión de la estrategia de operaciones desarrollada por B.F. Skinner, de Harvard. La sección 4 considera las cuestiones que surgen de la globalización de la manufactura e incluye una nueva Aplicación instantánea sobre la alianza estratégica de Read Rite's Corporation con una compañía de metales japonesa. Las secciones 5 y 8 consideran varias iniciativas estratégicas modernas llevadas a cabo en los años recientes. Éstas incluyen reingeniería de procesos de negocio, producción justo a tiempo, énfasis en la calidad y competencia basada en tiempo. La sección 9 considera el ciclo de vida del producto y cómo es necesario acoplar a los ciclos de vida del proceso y del producto.

La sección 10 se refiere a las curvas de experiencia y de aprendizaje. Éstas son ciertamente de importancia estratégica para cualquiera que esté interesado en proyectar costos y eficiencias. La sección final del capítulo 1 refiere modelos para planear la capacidad de crecimiento. Esta sección incluye también una discusión cualitativa de algunas cuestiones que surgen en las decisiones de ubicación de la planta (en el capítulo 10 se exploran a profundidad los modelos matemáticos usados para tomar la decisión de ubicación de una planta).

Capítulo 2. Pronósticos

Este capítulo trata a profundidad este importante tema. Se ha cuidado tanto el rigor como la accesibilidad del tema al público general. En este capítulo se revisan cuestiones como los periodos, desde y hacia dónde se pronostica.

Las primeras cuatro secciones del capítulo incluyen una discusión de varios temas importantes sobre pronósticos. En las secciones 5 y 6 se presenta la notación utilizada a lo largo de todo el capítulo y se definen las medidas de desempeño. La sección 7 está dedicada a discutir los métodos más populares para predecir series estacionarias, como los promedios móviles simples y el suavizamiento exponencial simple. Se aclara la diferencia entre los pronósticos de múltiples etapas y de una sola etapa. También se hace una cuidadosa comparación entre suavizamiento exponencial y promedios móviles. Se analizan los valores consistentes de N para

promedios móviles y α para el suavizamiento exponencial en el apéndice A, y probamos que se pueden obtener valores consistentes para estos parámetros a partir de la varianza de los errores del pronóstico. La Aplicación instantánea al final de la sección 7 habla sobre el sofisticado sistema de pronósticos de Sport Obermeyer.

La sección 8 considera métodos para pronosticar series en las que está presente una tendencia lineal. Consideramos regresión lineal simple y el método de Holt. Indudablemente, hay muchos otros métodos que podemos considerar, pero éstos parecen ser los más representativos. El método de Holt se presenta en vez del suavizamiento exponencial doble de R. G. Brown ya que, en opinión de este escritor, el método de Holt es más intuitivo.

La sección 9 aborda métodos para series estacionales. Presentamos un método popular para desestacionalizar una serie usando promedios móviles y una cuidadosa y detallada exposición del método de suavizamiento de Winters para series estacionales. Son pocos los textos en este nivel que consideran el método de Winters. La sección 10 considera varias cuestiones prácticas al momento de implementar los métodos de pronóstico. La sección 11 proporciona un breve panorama general de los métodos avanzados, incluyendo el pronóstico que contempla ventas perdidas. La sección 12 habla sobre la diferencia entre la varianza de la demanda y la varianza del error del pronóstico. Esta cuestión se trata en el capítulo 5.

Capítulo 3. Planeación agregada

Este capítulo presenta un tratamiento total de la planeación agregada y considera varias cuestiones que generalmente se pasan por alto. Tal es el caso del significado de una unidad agregada de producción. Los estudiantes muchas veces se sienten confundidos ante la diferencia entre las unidades normales y las unidades agregadas; muy pocos autores explican esta diferencia. Antes de presentar la metodología de la planeación agregada, comentamos la necesidad de su existencia y los costos relevantes, y proponemos un problema tipo caso que utilizaremos a lo largo de todo el capítulo.

La sección 4 considera los dos planes agregados más sencillos: el plan de fuerza laboral constante y el plan de inventario cero. La aplicación de la programación lineal para la planeación agregada se desarrolló con sumo cuidado. Debido a que la planeación agregada es un problema de programación entera, no se pueden ignorar las consideraciones que lo rodean. Presentamos un ejemplo detallado con salidas de computadora para ilustrar los pasos adecuados en torno a dicho problema.

Las secciones 7 y 8 consideran la Regla lineal de decisión de Holt, Modigliani, Muth y Simon, y los coeficientes gerenciales de Bowman. El material sobre disgregación de planes agregados de la sección 9 se toca en muy raras ocasiones dentro de los textos. La sección 10 considera cuestiones sobre la planeación agregada e incluye una aplicación instantánea.

Suplemento 1. Programación lineal

Este suplemento difiere de los capítulos únicamente en que no incluye problemas. El propósito de este suplemento es presentar rápidamente al estudiante la programación lineal, en particular para aquellos que probablemente no han estado expuestos a esta importante área. Conforme más programas académicos de negocios eliminan este tema de sus materias obligatorias en investigación de operaciones, los textos sobre administración de la producción y las operaciones tendrán que incluir material de apoyo para la enseñanza en estos cursos. Este suplemento es más detallado que la mayoría y está diseñado para brindar al estudiante suficientes antecedentes sobre la aplicación (a diferencia de la teoría) de la programación lineal como para poder implementarla. El principal cambio en esta sección es la eliminación de la presentación de LINDO. Se incluyó una amplia discusión sobre Solver, por ser ahora el estándar *de facto*.

Capítulo 4. Control de inventarios sujeto al conocimiento de la demanda

Los capítulos 4 y 5 tratan a profundidad los modelos de inventarios. Las secciones 1, 2 y 3 del capítulo 4 dan una descripción expositiva de los tipos de inventarios, la motivación para mantener

inventarios y las características de los sistemas de control de inventarios. La sección 4 presenta una discusión detallada de las formas de las funciones de costo en las que se basan los modelos de los capítulos 4 y 5. El tratamiento del modelo CEP simple en la sección 5 incluye discusiones sobre análisis de sensibilidad y tiempos de demora. La sección 6 versa sobre la extensión de la CEP simple para el caso de la tasa finita de producción.

La sección 7 se refiere a los modelos de descuento por cantidad incluyendo tanto descuentos para todas las unidades como los incrementales. El material de la sección 8, sistemas de múltiples productos con restricciones de recursos, rara vez se puede encontrar en textos de administración de operaciones. En esta sección mostramos bajo qué circunstancias pueden resolverse dichos problemas con facilidad y bajo qué condiciones se requieren los multiplicadores de Lagrange. La sección 9 sobre los modelos CEP para la planeación de la producción aborda la interfase entre la programación de la producción y los modelos de inventarios clásicos. La sección 10 da un breve resumen sobre software para el control de inventarios. La aplicación instantánea resalta el exitoso sistema de control de inventarios de Mervyn.

Capítulo 5. Control de inventarios sujetos a demanda incierta

Este capítulo presenta un tratamiento amplio y detallado de los modelos estocásticos de inventarios. Comienza con una discusión general de la aleatoriedad en el contexto del control de inventarios, al que le sigue un caso que sirve para ilustrar las metodologías presentadas en el resto del capítulo. Se incluyó una rigurosa derivación del modelo del repartidor de periódicos (o *newsboy*, por su nombre en inglés) así como se discuten con cuidado la interpretación de los costos. Elegí la formulación basada en los costos de excedente de mercancías y de faltantes de mercancías en lugar de la formulación basada en los costos de pedir, mantener y penalizar, que se encuentra en muchos textos de investigación de operaciones. Me he dado cuenta de que el estudiante puede comprender y aplicar mejor la primera formulación. Una característica única de nuestra cobertura es la extensión del modelo del repartidor a múltiples periodos de planeación.

La sección 4 es un tratamiento amplio de los modelos de tamaño del lote-punto de reorden. Los modelos de este tipo son la base de muchos de los sistemas comerciales para control de inventarios. La sección comienza con definiciones y descripciones de las variables y los costos. La sección 5 comprende niveles de servicio y la importante diferencia entre determinar la probabilidad de que no haya faltantes durante el tiempo de demora y la tasa de cumplimiento (o *fill rate*). Otras cuestiones tratadas en esta sección incluyen la escala de la demanda durante el tiempo de demora, el costo implícito de faltantes, la relación entre pronósticos e inventarios, y los tiempos de demora aleatorios.

Aunque una derivación normal de las políticas (s, S) está más allá del ámbito de este texto, las políticas (s, S) se mencionan junto con una breve discusión de los niveles de servicio en los sistemas con revisión periódica. La sección 7 considera el análisis ABC y las curvas de intercambio, y los sistemas de inventarios multiproducto. El capítulo concluye con un resumen de varios temas avanzados del modelado de inventarios y una breve discusión del intercambio electrónico de datos, IED (que se ve con mayor detalle en el capítulo 12).

Existen cinco apéndices al capítulo 5. Recomiendo que los profesores estén conscientes del material contenido en ellos porque es posible que deseen incluirlo en sus clases. El apéndice 5-A es una revisión básica de probabilidad. El apéndice 5-B agrega profundidad técnica considerable a la discusión del modelo del repartidor de periódicos, incluyendo la correcta interpretación de los costos de excedente y de faltante para diferentes versiones del horizonte infinito de pedidos pospuestos y los casos de ventas perdidas. El apéndice 5-C proporciona detalles de la derivación de la política (Q, R). El apéndice 5-D extiende los modelos tratados en el capítulo a las distribuciones de Poisson y de Laplace.

Capítulo 6. Administración de la cadena de abastecimiento

Este capítulo, nuevo en la cuarta edición, proporciona un panorama general completo de los temas clave en esta importante nueva área de la administración de operaciones. El capítulo comienza con definiciones y una aplicación instantánea sobre Wal-Mart. Las secciones 1-4 del capítulo

consideran el problema de transporte, y generalizaciones y formulaciones de redes más generales de los problemas de transbordo. La utilización de Excel para resolver estos problemas también se trata en esta sección. La sección 4 concluye con una aplicación instantánea del sistema de administración de refacciones de IBM.

La sección 5 se refiere al enfoque DRP para la planeación de recursos. La sección 6 sobre ruteo de vehículos aparecía anteriormente en el capítulo 7 de Programación de operaciones. La sección 7 trata sobre el diseño de producto para cadenas de abastecimiento; incluye material que anteriormente aparecía en el capítulo 12 de la tercera edición y algo de material nuevo. La Aplicación instantánea de esta sección es sobre computadoras Dell.

La sección 8 sobre el papel de la información en las cadenas de abastecimiento incluye material que ya aparecía en la tercera edición, así como nuevo material. También contiene una discusión sobre el efecto látigo. Las secciones 9 sobre sistemas de distribución multinivel y 10 sobre cadenas de abastecimiento global son nuevas. En la sección 9 hablo sobre algunos modelos de inventario multinivel y la forma en que se relacionan con tópicos de las modernas cadenas de abastecimiento. En la sección 10 abordo algunas cuestiones importantes que deben tomarse en cuenta al momento de diseñar una cadena de abastecimiento. El capítulo concluye con otra aplicación instantánea en la que se trata el diseño de la cadena de abastecimiento de DEC.

Capítulo 7. Sistemas de producción de empuje y tracción: MRP y JIT

Tratamos a MRP y JIT como dos filosofías para el control de manufactura. La primera mitad de este capítulo se refiere a MRP y al dimensionamiento de lotes. Este capítulo tiene varias características únicas. Una es que el dimensionamiento dinámico de lotes se incluye en este punto, en lugar de ser estudiado en el capítulo 4 o en un capítulo separado sobre planeación de la producción. De esta forma se deja clara la importancia del dimensionamiento de lotes para MRP. La manera en que el dimensionamiento de lotes para un determinado nivel afecta al patrón de los requerimientos de los niveles inferiores es algo que se pasa por alto con frecuencia al exponer dicho tema.

Se menciona un estudio de caso a fin de mostrar la forma en que funciona el cálculo de la explosión. El ejemplo se realiza para tres niveles de la estructura de un producto. El dimensionamiento óptimo de lotes utilizando el algoritmo de Wagner-Whitin ha sido relegado al apéndice A y puede pasarse por alto sin perder continuidad. Consideramos también las heurísticas de Silver Meal, del mínimo costo unitario, y de balanceo de partes y periodos. (La heurística de la cantidad periódica de pedidos se comenta también en la parte de problemas.) Al introducir aquí estos algoritmos de dimensionamiento de lotes, podemos mostrar la forma en que las decisiones que los involucran impactan al plan de requerimientos de materiales de los diferentes niveles de la estructura de un producto. También, presentamos una heurística original (similar a las otras que han aparecido en la literatura) para el dimensionamiento de lotes bajo restricciones.

La sección 7 proporciona un tratamiento a profundidad del justo a tiempo (JIT). Parte de este material apareció en el capítulo 12 de las ediciones previas, pero también hay varios temas nuevos, incluyendo Intercambio de dados de un minuto (SMED) y Trabajo en proceso constante (CONWIP). Se comparan y contrastan las filosofías MRP y JIT. El capítulo concluye con una discusión sobre el creciente mercado para el software de MRP.

En la sección 8 de este capítulo se habla sobre problemas que surgen al momento de implementar un sistema MRP. Éstos incluyen la incertidumbre, la planeación de la capacidad, los horizontes rodantes y la nerviosidad del sistema, por mencionar algunos.

Capítulo 8. Programación de operaciones

Este capítulo presenta un panorama general de los resultados más importantes para la programación de una secuencia. El efecto de las reglas de secuenciación en varias medidas de desempeño queda claro al presentar un ejemplo detallado en el que se muestra el efecto de las cuatro reglas de secuenciación: primeras llegadas, primeras salidas; tiempo de procesamiento más corto primero; primera fecha de entrega, y razón crítica. Se presentan varios algoritmos para la

secuenciación en una máquina, incluyendo el algoritmo de Moore para minimizar el número de trabajos tardíos y el algoritmo de Lawler para restricciones de precedencia.

La siguiente porción del capítulo considera los algoritmos de secuenciación para múltiples máquinas. Presentamos el algoritmo de Johnson para la programación de dos y tres máquinas y el procedimiento de Aaker para dos trabajos en *m* máquinas. Éste es quizás el único texto que considera también resultados para programación estocástica. Exponemos varios de los resultados más elegantes en esta área e incluimos algunas pruebas para mejorar la comprensión del material por parte del estudiante. En esta edición también se discute la programación estocástica en un ambiente dinámico. Esta sección es una aplicación de la teoría de colas para la programación de una secuencia.

El último tema importante a tratar en el capítulo 8 es el balanceo de una línea de ensamble. Presentamos sólo una heurística para resolver ese problema, la técnica de peso posicional clasificado. La sección trata sobre la experimentación con diferentes valores del tiempo de ciclo a fin de lograr un balance óptimo.

Se proporciona aquí una breve revisión de software post-MRP para programación, así como un panorama general para usar la simulación al calendarizar. La aplicación instantánea proporciona descripciones de varias iniciativas de programación exitosas en el mundo real.

Suplemento 2. Teoría de colas

Este material apoya la sección sobre programación estocástica dinámica del capítulo 8 al proporcionar al estudiante un buen panorama general de la teoría de colas. Incluye una derivación de la probabilidad para el estado estable para la cola M/M/1 y diferentes resultados, sin demostrarlos, para modelos más complejos. Al igual que en el caso del suplemento de Programación lineal, puede resultar de ayuda para los estudiantes que anteriormente no han estado expuestos a la teoría de colas.

Capítulo 9. Programación de proyectos

Este capítulo comienza con una descripción del problema de programación de proyectos y la forma en que se representan los proyectos como redes. Se introduce un estudio de caso para ilustrar las metodologías comentadas a lo largo del resto del capítulo. Presentamos los detalles del método de la ruta crítica (CPM), incluyendo el cálculo de los tiempos de inicio y terminación más próximos y más tardíos, así como la identificación de la ruta crítica. En la sección 4 se consideran también métodos de costeo.

El método PERT para el manejo de tiempos aleatorios de las actividades se trata en la sección 5. Se presenta el enfoque PERT tradicional, que asume que la longitud de la ruta crítica tiene una distribución normal y es la ruta con el tiempo de actividad esperado más largo. Una aproximación alternativa consiste en suponer que el tiempo requerido para completar dos o más rutas es una variable aleatoria independiente. En algunos casos, la suposición de la independencia de rutas puede dar resultados más exactos que el enfoque PERT tradicional, que asume que la ruta con el valor esperado más largo equivale a la ruta más larga.

Este capítulo incluye material raramente cubierto en los textos de producción. Mostramos un ejemplo sobre la forma en que se incorporan restricciones de recursos en el marco de trabajo típico de la planeación de proyectos y cómo construir perfiles de carga de recursos. Este capítulo incluye una sección que muestra cómo la programación lineal puede aplicarse para resolver problemas de programación de proyectos; con el mismo objetivo también puede utilizarse software especializado, el cual se revisa en este punto. Se analizan también las estructuras organizacionales que facilitan la administración de proyectos, así como software para la administración de proyectos en una computadora personal.

Capítulo 10. Distribución y ubicación de instalaciones

El capítulo 10 trata de manera global los principales desarrollos en el área de ubicación y disposición. Se usan ejemplos abundantes para ilustrar una metodología compleja. Examinamos los tradicionales diagramas que ayudan a tomar decisiones de disposición, incluyendo el diagrama

desde-hacia y el diagrama rel. En la sección 4 mostramos la forma en que se usa el modelo de asignación simple para resolver los problemas de ubicación en los que no hay interacción entre las instalaciones, y un conjunto pequeño de ubicaciones alternativas. Asimismo, se presenta el modelo cuadrático de asignación para resolver problemas más complejos, aunque los detalles de implementación del algoritmo van más allá del alcance de este texto.

La sección 6 da una descripción detallada de CRAFT. Mostramos la forma en que se calculan los centroides (usados por CRAFT) en el apéndice A y aplicamos el método en un ejemplo particular. Otros métodos de disposición computarizados se mencionan con menor detalle: éstos incluyen COFAD, ALDEP, CORELAP y PLANET. Esta sección concluye con un interesante debate proveniente de la literatura especializada, acerca de quiénes realizan mejores disposiciones, si las computadoras o los seres humanos.

En las secciones 8 a 11 del capítulo 10 tratamos el problema de ubicar nuevas instalaciones. Consideramos tópicos cualitativos de la ubicación tratados en detalle en el capítulo 1 como parte de nuestra presentación sobre estrategia, de manera que la discusión no se repite aquí. En la sección 9 hablamos del método para ubicar una nueva instalación sujeta a una medida de distancia rectilínea. Otros temas comprendidos en esta sección incluyen curvas de nivel y un objetivo minimax, en lugar de un objetivo ponderado. En la sección 10 se presentan las medidas de distancia euclidiana y del cuadrado de la distancia euclidiana. En la sección 11 incluimos discusiones breves sobre otras extensiones, entre las que podemos mencionar modelos para ubicar múltiples instalaciones, problemas de ubicación-asignación, modelos de ubicación discreta y modelos de ubicación en redes.

Capítulo 11. Aseguramiento de la calidad

Este capítulo es uno de los más grandes en el texto y cubre a profundidad las cuestiones y técnicas relevantes del área de la calidad. La primera parte del capítulo está dedicada a las gráficas de control. Primero se revisan brevemente los resultados relevantes de la teoría de probabilidad. Igual que en la mayoría de las técnicas comentadas en el texto, comenzamos la sección con un ejemplo de tipo caso. Este caso se usa para ilustrar la construcción de las gráficas $\overline{X}$ y R. Un punto comentado en la sección 2 que otras presentaciones muchas veces pasan por alto se refiere a la interpretación de los errores Tipo 1 y Tipo 2 en el contexto de las gráficas de control. También se abordan las gráficas p y c.

La sección 6 incluye material que se encuentra muy pocas veces en textos de tipo panorámico, a saber, el diseño económico de las gráficas $\overline{X}$. Asumiendo que se conocen los costos por muestreo, de búsqueda de una causa asignable y de operación fuera de control, mostramos cómo encontrar el diseño de mínimo costo esperado de la gráfica $\overline{X}$. La teoría se ilustra con uno de los ejemplos de caso del capítulo.

La segunda parte del capítulo está dedicada al muestreo de aceptación. De nueva cuenta se introdujo un caso como ejemplo para ilustrar la metodología. Comentamos algunos planes de muestreo simple, de muestreo doble y de muestreo secuencial. También incluimos en esta sección un tratamiento detallado de la calidad saliente promedio con derivaciones de las diferentes fórmulas. El capítulo concluye con una discusión sobre diferentes métodos populares de administración para mejorar la calidad incluyendo círculos de calidad, control total de la calidad y cero defectos.

La sección 13 proporciona una exposición de la administración de la calidad total incluyendo definiciones, la competencia basada en la calidad, la organización para la calidad, calidad benchmarking, y una discusión sobre el Premio Deming Prize y el Premio Baldrige. También se incluyen secciones sobre cómo escuchar al cliente y el despliegue de la función de calidad. La sección 14 examina la interfase entre la calidad y el diseño y la aplicación de los métodos de Taguchi. También se incluye material sobre diseño para manufacturabilidad que estaba en el capítulo 12 de la tercera edición.

Capítulo 12. Confiabilidad y facilidad de mantenimiento

Aunque la confiabilidad y el mantenimiento se consideran parte del cuerpo de información comprendido en la administración de operaciones y la producción, pocos textos con visión panorá-

mica incluyen este material. En parte, se debe a que, matemáticamente hablando, la confiabilidad es un tema difícil. Sin embargo, la confiabilidad y la calidad están muy unidas: parte del control de calidad reside en comprender cómo fallan las partes. Es más, muy pocas personas se opondrían a la aseveración de que el mantenimiento es una parte importante del análisis de la producción y las operaciones, y los modelos de mantenimiento se basan en la teoría de la confiabilidad.

Debido a la naturaleza del material, este capítulo es quizás el de mayor sofisticación matemática en el texto. Se han realizado todos los intentos posibles para que el texto resulte lo más disfrutable posible sin trivializarlo. Se han utilizado muchos ejemplos para propiciar en el lector la intuición de conceptos difíciles como la función de la tasa de fallas y la aleatoriedad pura asociada con la ley exponencial de fallas.

El proceso de Poisson es una parte importante de este capítulo. Lo presentamos cuando hablamos de sistemas en serie de componentes que fallan sólo de manera aleatoria, pero también lo utilizamos al final del capítulo. La porción de confiabilidad del capítulo concluye con una discusión de las leyes de falla para grupos de componentes.

Las secciones 5 a 7 del capítulo 12 están dedicadas a los modelos de mantenimiento. Tratamos modelos de reemplazo por edad tanto para procesos con fallas determinísticas como aleatorias. Desarrollamos un modelo determinístico general usando funciones exponenciales para la depreciación y para los costos. El reemplazo planeado bajo incertidumbre para un artículo y para un bloque de artículos es el tema tratado en la sección 7.

En la sección 8 consideramos garantías. Dicho tema se menciona en muy pocas ocasiones en los libros de texto. La sección se basa en resultados de la literatura especializada. Se utilizan ejemplos de la vida real para ilustrar el valor práctico de esta teoría. El capítulo concluye con un breve panorama general de las cuestiones clave de la confiabilidad de software.

REFERENCIAS

Bowman, E. H. y R. B. Fetter. *Analysis for Production and Operations Management.* 3a. ed., Nueva York: McGraw-Hill/Irwin, 1967.

Skinner, W. “The Taming of Lions: How Manufacturing Leadership Evolved, 1780-1984”, Capítulo 2 en *The Uneasy Alliance: Managing the Productivity-Technology Dilemma,* editado por K. B. Clark, R. H. Hayes y C. Lorenz. Boston, MA: Harvard Business School Press, 1985.

Capítulo **Uno**

Estrategia y competencia

Panorama general del capítulo

Objetivo

El propósito de este capítulo es presentar al estudiante una variedad de temas estratégicos que surgen en la función de manufactura de una empresa.

Puntos clave

1. *La manufactura es relevante.* La pérdida de su base de manufactura no favorecerá la economía estadounidense; con el tiempo, causará un deterioro general en el nivel y calidad de vida. Esto se contrapone a la idea de que es benéfico evolucionar hacia una economía de servicio.
2. *Dimensiones estratégicas.* Junto con el costo y/o diferenciación de producto, existen otras características que distinguen a las empresas: *a*) la calidad, *b*) la velocidad de entrega, *c*) la confiabilidad en la entrega y *d*) la flexibilidad.
3. *Perspectiva clásica.* La literatura clásica en torno a estrategias de manufactura indica que la estrategia debería considerarse en relación con uno o más de los siguientes temas: *a*) horizonte de tiempo, *b*) focalización, *c*) evaluación y *d*) consistencia.
4. *Competencia mundial.* ¿Cómo medimos nuestro éxito y bienestar económicos en una escala mundial? Una de las maneras consiste en analizar las medidas clásicas de la fortaleza económica relativa, que incluye: *a*) balanza comercial, *b*) participación en las exportaciones mundiales, *c*) la creación de empleos y *d*) el costo de mano de obra. Sin embargo, tales macromedidas no explican adecuadamente por qué ciertos países dominan industrias específicas. La ventaja competitiva nacional es una consecuencia de varios factores (condiciones causales; condiciones de demanda; industrias relacionadas y de apoyo; estructura estratégica de la empresa; rivalidad), aunque la productividad también juega un papel importante.
5. *Iniciativas estratégicas.* Aquí hablaremos sobre varias iniciativas estratégicas, las cuales han permitido a muchas empresas destacar en sus campos respectivos. Entre tales iniciativas se incluyen: *a*) reingeniería de procesos de negocio, *b*) manufactura y sistemas de compra "justo a tiempo", *c*) competencia basada en el tiempo y *d*) competencia con la calidad.
6. *Ciclos de vida del producto y del proceso.* La mayoría de nosotros entendemos que los productos tienen ciclos de vida naturales: inicio, crecimiento rápido, maduración y estabilización o declive. No obstante, comúnmente no se reconoce que los procesos también tienen ciclos de vida. Al principio, los procesos de manufactura tienen las características de un "taller". Según madura el proceso, se da paso a la automatización. En las fases maduras de un proceso de manufactura, la mayoría de las operaciones

importantes son automáticas. Para ser exitosa, una empresa debe hacer coincidir las fases de los ciclos de vida de un producto y sus procesos.

7. *Curvas de aprendizaje y de experiencia.* Éstas son útiles para pronosticar la reducción en el costo unitario de un proceso de manufactura conforme se adquiere experiencia en el proceso. Las curvas de aprendizaje resultan más apropiadas cuando se modela el aprendizaje de un trabajador, y las de la experiencia cuando se considera toda una industria.
8. *Planeación del crecimiento de la capacidad.* Otro tema estratégico importante en las operaciones consiste en la determinación del tiempo oportuno y las dimensiones que ha de tener la expansión de la capacidad. Los modelos simples (el problema de fabricar o comprar) y algunos modelos más complejos de crecimiento exponencial se analizan en la sección 1.11. Además, se exploran algunos de los factores que determinan la ubicación apropiada de nuevas instalaciones.

El éxito requiere visión, y las visiones se deben articular de manera tal que los empleados de la empresa puedan compartirla. La articulación formal de la visión se conoce como declaración de los objetivos de la empresa. Una buena declaración de objetivos debe proporcionar una descripción clara de las metas de una empresa, y es el primer paso hacia la formulación de una estrategia comercial coherente. Las declaraciones de objetivos deficientes tienden a la verborrea y están llenas de generalidades. Una buena declaración de objetivos es directa, clara y concisa. Jones y Kahaner (1995) enumeran las que consideran las mejores 50 declaraciones corporativas de objetivos. Un ejemplo es el de Gillette Corporation. Su declaración de objetivos es: "Nuestra misión es lograr o aumentar un liderazgo claro, mundial, en las categorías ya existentes o en las nuevas categorías de productos de consumo primordiales en las que queramos competir". Después enumeran con exactitud cuáles áreas consideran clave. Intel define así su misión: "Realizar un excelente trabajo para nuestros clientes, empleados y accionistas al ser el proveedor fundamental predominante para la industria de la computación". Su declaración de objetivos enseguida proporciona detalles sobre "valores y objetivos". En muchos casos, sus objetivos son bastante específicos (por ejemplo, "Líder en productos LAN y Smart Network Services"). Desde luego, el premio por ser concisos se lo lleva General Electric Corporation, cuya declaración de objetivos se reduce a tres palabras: "Sin límites… Velocidad… Extensión". El comentario que sigue a la declaración de objetivos explica lo que realmente significan estas palabras en el contexto de una empresa.

Una vez que se ha articulado la visión, el siguiente paso consiste en planear la estrategia para lograrla. Ésta es la estrategia comercial de la empresa. En la estrategia comercial global se define:

1. El mercado en el que la empresa compite.
2. El nivel de inversión.
3. Los medios para ubicar recursos e integrar las diferentes unidades de negocio.
4. Estrategias de áreas funcionales, que incluyen:
   - La estrategia de mercadotecnia
   - La estrategia financiera
   - La estrategia de operaciones

A grandes rasgos, la estrategia de operaciones es la manera en que la empresa utiliza sus recursos para lograr sus metas competitivas. Para las empresas de manufactura, equivale al total de todas las decisiones relacionadas con la producción, almacenamiento y distribución de mercancías. Algunas decisiones importantes en la estrategia de operaciones incluyen la selección de la ubicación de nuevas instalaciones de manufactura, de sus dimensiones y de los procesos que se utilizarán para manufactura; el desplazamiento de bienes dentro del sistema, y a

qué trabajadores emplear. Las empresas de servicio también requieren una estrategia de operaciones. Estados Unidos sigue siendo un líder en los servicios financieros; para continuar siendo competitivos, se necesita un apoyo efectivo y confiable por parte de los departamentos de operaciones. El historial de éxitos de Disneylandia, en parte obedece al cuidado puesto en los detalles de cada etapa de sus operaciones.

¿La cultura estadounidense enfatiza demasiado la mercadotecnia (para vender el producto) y las finanzas (compras apalancadas, fusiones, precios de acciones), y atiende poco las operaciones (hacer y entregar el producto)? Esto era verdad hace muchos años. Sin embargo, aprendemos rápido. El enorme éxito de la industria automotriz japonesa, por ejemplo, motivó que las tres grandes empresas automotrices estadounidenses más importantes cerraran sus plantas ineficientes y cambiaran su manera de operar. Las grandes diferencias que Womack, Jones y Roos (1990) dieron a conocer se han eliminado en su mayoría. Hoy en día, las plantas automotrices estadounidenses compiten con sus contrapartes japonesas en cuanto a calidad y eficiencia.

Aun así, sigue siendo esencial una estrategia de operaciones coherente. Cuando se introdujo la computadora Apple Macintosh, el producto fue sumamente exitoso. Sin embargo, la empresa se vio asediada con pedidos pospuestos y no pudo satisfacer la demanda de los consumidores. De acuerdo con Debbi Coleman, ex directora de manufactura mundial de Apple:

> La manufactura carecía de una estrategia global, lo cual creaba problemas que tardaban nueve meses en ser resueltos… realizábamos pronósticos extremadamente malos. No inspeccionábamos la materia prima para asegurarnos de que no tuviera defectos y, salvo llamadas telefónicas llenas de furia, no teníamos un mecanismo para indicar a nuestros proveedores qué era lo que estaba saliendo mal. Cuarenta por ciento de la materia prima de Mac provenía del extranjero y no había nadie de Apple que la inspeccionara antes de su embarque… Una de las más importantes tareas que los fabricantes de alta tecnología enfrentan consiste en diseñar una estrategia de manufactura que proporcione a una compañía la flexibilidad suficiente para hacer frente a las altas y las bajas de los ciclos de compra de los consumidores y de las empresas (Fallon, 1985).

Aunque es muy fácil criticar el estilo de administración estadounidense, debemos tener en mente los factores que motivan a los gerentes estadounidenses, y los que motivan a los gerentes de otras culturas. Los japoneses, por ejemplo, no han logrado sus impresionantes éxitos a cambio de nada. En Japón son comunes los días laborales de 16 horas y una elevada tasa de crisis nerviosas entre los empleados con puestos directivos.

Medir el éxito de una empresa según el desempeño del precio de sus acciones puede dar como resultado prácticas administrativas con poca visión de futuro. A los consejos de administración les preocupa más el próximo reporte trimestral que financiar proyectos importantes a largo plazo. De hecho, Hayes y Wheelwright (1984) argumentan, de manera convincente, que tales factores originan en Estados Unidos una miopía administrativa, que se caracteriza por:

1. El desempeño de los gerentes se mide con base en el **rendimiento sobre la inversión** (en inglés ROI o *return on investment*), que es simplemente la razón entre la ganancia obtenida por una operación o proyecto en particular y la inversión realizada para esa operación o proyecto.
2. El desempeño se mide según horizontes de tiempo cortos. Existe poca motivación para que un gerente invierta en un proyecto que probablemente dé frutos hasta después de que dicha persona se haya cambiado a otro puesto.

A fin de mejorar el ROI, un gerente debe o aumentar el numerador (ganancias) o disminuir el denominador (inversión). A corto plazo, es más fácil disminuir el denominador reduciendo la inversión en nuevas tecnologías o en nuevas instalaciones, que tratar de aumentar las ganancias mejorando la eficiencia, la calidad del producto o la productividad de la unidad de operación. Los efectos a largo plazo de la disminución de inversión son devastadores. En algún momento, los costos de capital requeridos para modernizar fábricas viejas pueden ser más de lo que la empresa puede financiar, con lo que pierde competitividad en el mercado.

Sería alentador que se estuvieran reduciendo los problemas de las industrias causados por un énfasis excesivo en el desempeño financiero a corto plazo, pero tristemente, parecen empeo-

rar. Debido a una pésima administración y a dudosas prácticas de auditoría, en 2001 cayeron dos gigantes de la industria estadounidense: Enron y Arthur Andersen. "De ser la empresa en séptimo lugar de la lista en *Fortune 500*, las acciones de Enron cayeron a centavos en sólo tres semanas; aparentemente, mintió en sus informes financieros", dijo el representante John D. Dingell, quien perteneció al Comité de Energía de la Cámara de Representantes del Congreso de Estados Unidos. Mientras otras partes del mundo también han experimentado problemas espectaculares (como la crisis financiera asiática que ocurrió a finales de los años noventa), pocos estadounidenses pueden entender cómo esa empresa que recientemente había aumentado en tamaño y que se benefició de la crisis de energía, un ícono estadounidense como Arthur Andersen, pudiera derribarse tan rápido y por completo. Nuestro continuo énfasis centrado en el desempeño a corto plazo y el sistema de incentivos construido orientado a alcanzar sólo este tipo de objetivos producen esta clase de crisis.

Medir el desempeño individual durante un corto plazo es una filosofía que parece extenderse a toda la vida estadounidense. Se elige a los políticos para mandatos de dos, cuatro o seis años. Los alentamos a dar resultados a tiempo para las siguientes elecciones. Hasta los profesores universitarios, en muchas instituciones, son evaluados anualmente en relación con su desempeño profesional, aun cuando la mayoría de los proyectos académicos importantes requieren de muchos años para su maduración.

1.1 LA MANUFACTURA ES RELEVANTE

Una cuestión que está siendo debatida por los economistas durante los últimos 20 años es la importancia que tiene una sólida base de manufactura. El declive de la manufactura a nivel nacional ha provocado un desplazamiento laboral del sector de manufactura al de servicios. Debido a las disparidades significativas en los costos de la mano de obra en las diferentes partes del mundo, las empresas estadounidenses se ven fuertemente motivadas a ubicar sus instalaciones de manufactura de gran escala en el extranjero, para reducir los costos de mano de obra. ¿Es importante una fuerte base de manufactura para el bienestar de la economía?

Una gran parte de los bienes manufacturados proviene del extranjero. Dertouzes y otros (1989) utilizan información del Departamento de Comercio para mostrar que, para el periodo de 1972 a 1986, las importaciones en ocho de los sectores industriales más importantes presentaron un incremento significativo en su partipación porcentual del mercado global estadounidense, mientras que las exportaciones de estos mismos sectores se redujeron significativamente. [Estos sectores son: automtriz, químico, aeronaves comerciales, electrónica para el consumidor, máquinas herramientas, semiconductores (incluyendo computadoras y equipo de oficina), acero y textiles.] Esto demuestra que la manufactura en estas industrias se trasladó fuera del país a un ritmo veloz. ¿Debe esto preocuparnos?

Personalidades académicas (por ejemplo, Daniel Bell, 1976) argumentan que esto representa sólo la evolución natural de una economía industrial hacia una de servicios. Desde esta perspectiva, las tres etapas de evolución económica serían: 1) agraria, 2) industrial y 3) de servicios. En los primeros años de nuestro país teníamos una economía primordialmente agraria. Con la revolución industrial, una gran parte de la fuerza laboral pasó de la agricultura a la manufactura. En los últimos años parece haber menos interés en esta última. Dichos autores sostienen que sencillamente estamos entrando a la tercera etapa del proceso de evolución: pasar de la economía industrial a la de servicios.

Es reconfortante pensar que la economía estadounidense es saludable y que sólo evoluciona de una economía industrial hacia una de servicios. Uno también podría aducir que la manufactura no es importante para el bienestar económico. De acuerdo con el economista Gary S. Becker (1986), "las economías modernas fuertes no parecen requerir un sector de manufactura dominante".

Sin embargo, no queda suficientemente claro que hayamos evolucionado de una economía agraria hacia una economía industrial. Aunque menos trabajadores estadounidenses son emplea-

dos en el sector agrícola de la economía, la producción agrícola *no* ha menguado. Con base en datos del Departamento de Comercio de Estados Unidos, Cohen y Zysman (1987) dicen que "la agricultura ha mantenido, durante un largo periodo, la tasa más alta de aumento de productividad de entre todos los sectores". Al utilizar nuevas tecnologías, la agricultura ha podido mantener su crecimiento, utilizando menos horas de mano de obra. Por lo tanto, las cifras no confirman que nuestra economía haya cambiado de una economía agrícola a una economía industrial.

Los hechos no permiten afirmar que la economía esté experimentando etapas naturales de evolución. Considero que todos los sectores de la economía —agrícola, de manufactura y de servicios— son importantes, y que el bienestar de la economía doméstica depende de que las actividades de estos sectores se vinculen correctamente.

El rendimiento de las innovaciones se perderá si los nuevos productos se abandonan después del desarrollo. Los frutos de la investigación y desarrollo (IyD) sólo pueden recogerse una vez que un producto se produce y se vende. Si la manufactura se realiza en el extranjero, no puede recuperarse "el rendimiento sobre la innovación". Además, la manufactura de manera natural lleva a la innovación. Para Estados Unidos sería difícil mantener su posición como líder en innovación si pierde su posición como líder en la manufactura.

El que la manufactura conduce de manera natural a la innovación puede ejemplificarse mejor con la experiencia japonesa en el mercado de video. Después de que Japón captó la parte más grande del mercado mundial de televisiones, la siguiente gran innovación en tecnología de video para el consumidor, la videograbadora (VCR) (por lo menos, la versión económica) fue desarrollada en Japón, no en Estados Unidos. Prácticamente todas las videograbadoras que se venden actualmente se fabrican en Asia.

Es difícil sostener el argumento de que el paso de una economía industrial a una de servicios puede resultar sencillo. Existen muchos servicios que se dedican a apoyar la manufactura. Si las actividades de manufactura se cambian al extranjero, es muy seguro que sufrirán los servicios que la complementan. Según Cohen y Zysman (1987):

> Si Estados Unidos deja de dominar y controlar su manufactura, muchos de los trabajos de servicio se irán reubicando después de varios ciclos de innovación de los productos y los procesos. Básicamente, saldrán de Estados Unidos, y los salarios reales en todas las actividades de servicio caerán, empobreciendo a la nación.

Sin embargo, hay noticias positivas. Los costos unitarios de mano de obra en el sector de manufactura de Estados Unidos apenas han crecido durante los últimos años (a una tasa anual promedio de alrededor de 1%), mientras que la productividad laboral de la manufactura de Estados Unidos ha aumentado a una tasa promedio anual de alrededor de 3% (Agencia de Estadísticas Laborales de Estados Unidos). Estados Unidos todavía posee la mayor parte de la manufactura mundial. Sin embargo, la balanza de pagos de Estados Unidos todavía es sombría. El balance en la cuenta para 2004 se proyecta en números rojos de 480 millardos de dólares —más del triple de las cifras en 1994 (*World Economic Outlook*, marzo de 2003).

Si los empleos de manufactura se están trasladando al extranjero, ¿habrá menos oportunidades para nuevas carreras en la manufactura en los siguientes años? Según el *Manual de la perspectiva ocupacional para 2002-2003* publicado por el Departamento de Comercio,[1] se espera que el empleo en manufactura en Estados Unidos disminuya durante los siguientes años. Esta es una tendencia que no creemos que cambie en el futuro próximo. Sin embargo, para entonces, la composición del empleo de la manufactura irá de empleos relacionados con la producción a puestos de tipo profesional y gerencial. Por lo tanto, habrá una mayor demanda de profesionales con la capacitación adecuada para administrar la función de manufactura. Los puestos gerenciales para manufactura y los empleos relacionados con el desarrollo de nuevas tecnologías de manufactura continuarán creciendo.

---

[1] Disponible en http://www.bls.gov/oco/

La innovación es un medio por el cual algunas compañías estadounidenses han logrado mantener la delantera frente a sus competidores. Algunos ejemplos importantes son Sun Microsystems en California y 3M Company en Minnesota. Sun subcontrata una parte importante de su manufactura y ha mantenido una impresionante participación del lucrativo mercado de estaciones de trabajo mediante la introducción innovadora y oportuna de nuevos productos. 3M opera con la filosofía de que 70% de sus ventas en los próximos cinco años provendrá de productos que no existen todavía (Drucker, 1991).

Sin embargo, preocupantes tendencias sugieren que los japoneses, en particular, pueden estar sobrepasando a los estadounidenses en innovaciones. En 1980, los mayores 10 ganadores de patentes en Estados Unidos (en orden) fueron General Electric, Bayer, RCA, la Marina de Estados Unidos, AT&T, IBM, Hitachi, Westinghouse Electric, Siemens y General Motors. En 1990, la lista fue encabezada por Hitachi, Toshiba, Canon, Mitsubishi, General Electric, Fuji Foto, Eastman Kodak, Philips, IBM y Siemens. No se trata únicamente de que la mitad de las compañías de esta lista sean japonesas, sino que incluyen a cinco de las primeras seis ganadoras de patentes (Stewart, 1991). Datos más recientes muestran, no obstante, que Estados Unidos continúa ocupando el primer lugar en el número de patentes emitidas a nivel mundial.

1.2 UN MARCO DE TRABAJO PARA LA ESTRATEGIA DE OPERACIONES

La literatura clásica sobre competitividad establece que una compañía se puede posicionar estratégicamente en el mercado a través de una de dos dimensiones: costo más bajo o diferenciación del producto (Porter, 1990).

Con frecuencia, los nuevos participantes en un mercado se suelen posicionar como proveedores de bajo costo. Entre las compañías que han adoptado este enfoque están los fabricantes de autos coreanos (Hyundai, Daewoo, Kia), tiendas de descuento como Costco y minoristas como Wal-Mart. Aunque un proveedor de bajo costo puede tener éxito a corto plazo, corre riesgos. Los consumidores finalmente abandonarán los productos que consideren de baja calidad, sin importar cuál sea el costo. Por ejemplo, muchos fabricantes de clones de computadoras personales de bajo costo, que fueron muy populares en los ochenta, desaparecieron desde hace mucho tiempo.

La mayoría de las firmas que han tenido éxito en el mercado a lo largo del tiempo han logrado diferenciarse de sus competidores. Al proporcionar calidad única a los compradores, pueden sostener márgenes de utilidades a largo plazo. Un ejemplo es BMW, una de las firmas de automóviles más rentables del mundo. BMW continúa produciendo autos muy bien hechos de alto desempeño, que con frecuencia son sustancialmente más caros que los fabricados por la competencia en su ramo. La diferenciación de productos dentro de una misma compañía también ha probado ser una estrategia exitosa. Compárese el éxito de General Motors en los primeros años con el obtenido por Ford. GM fue capaz de captar diferentes segmentos de mercado al mismo tiempo mediante la formación de cinco divisiones diferentes; mientras que al insistir en proporcionar un modelo único, Henry Ford casi quiebra (Womack *et al.*, 1990).

Dimensiones estratégicas

Sin embargo, el costo y la diferenciación de producto no son las únicas dimensiones a través de las cuales las compañías pueden diferenciarse. Los siguientes factores se relacionan directamente con la función de operaciones:

- Calidad
- Velocidad de entrega
- Confiabilidad en la entrega
- Flexibilidad

¿Qué significa *calidad*? Es una palabra que se esgrime con frecuencia, pero significa cosas distintas en diferentes contextos. Considérense los siguientes comentarios.

1. “Esa secadora fue una verdadera desilusión. La verdad no me seca el pelo como yo lo imaginaba.”

2. "Estaba encantado con mi último coche. Lo vendí cuando tenía 150 000 millas y prácticamente nunca necesitó ninguna reparación."
3. "Me fascina comprar en ese catálogo. Siempre recibo la mercancía antes de 48 horas."
4. "El refrigerador funciona bien, pero creo que sus cajones y anaqueles podrían estar mejor ubicados."
5. "Ese parque tenía muy buenos juegos mecánicos. Pero las filas de espera eran todo un lío."
6. "Nuestra calidad es maravillosa. Tenemos menos de seis defectos por millón de partes producidas."

En cada situación, el interlocutor se refiere a un aspecto distinto de la calidad. En el primer caso, el producto sencillamente no cumple con la tarea para la cual fue diseñado. Esto es, su funcionamiento era de baja calidad. El historial de las reparaciones de un automóvil es más una cuestión de confiabilidad que de calidad, *per se*. En el tercer caso, la velocidad de entrega se traduce en un servicio de calidad para el cliente. El cuarto caso se refiere a un producto que hace lo que se supone que debe hacer, pero el cliente se siente desilusionado con el diseño del producto. La calidad del producto (los juegos mecánicos) en el parque de diversiones estaba bien, pero la logística de la administración del parque era desilusionante. El caso final se refiere a aspectos estadísticos del control de la calidad.

En efecto, la palabra *calidad* significa distintas cosas en diferentes contextos. Un Geo Prism es un producto de calidad y lo mismo pasa con un Ferrari Testarosa. Los consumidores que adquieren estos productos buscan autos de calidad pero tienen objetivos fundamentalmente diferentes. El hecho es que *todo mundo* compite en términos de calidad. Por esta razón, Terry Hill (1993) clasifica la calidad como algo que permite calificar los pedidos de los consumidores, más que algo que garantiza el ganar pedidos. Una opción queda inmediatamente descartada si no cumple con estándares de calidad mínimos. El aspecto específico de calidad que se ha seleccionado determina la estrategia competitiva y el posicionamiento de la compañía.

Velocidad de entrega puede ser un arma competitiva importante en ciertos contextos. Algunas compañías basan su competencia en la velocidad de entrega, como UPS y Federal Express. Los minoristas por correspondencia o vía Internet también deben ser capaces de entregar productos de forma confiable y rápida para mantener su competitividad. Los contratistas del campo de la construcción que terminan proyectos a tiempo aventajarán a los que no.

Confiabilidad de la entrega significa ser capaz de entregar productos o prestar servicios en el tiempo prometido. Los corretajes en línea que negocian rápida y confiablemente conservarán a sus clientes. Los fabricantes contratistas se evalúan de varias maneras, una de ellas es su capacidad de entrega oportuna. Conforme continúa creciendo la contratación externa de manufactura, los fabricantes contratistas exitosos serán los que ponen primero a los clientes y conservan un historial de entrega de productos de alta calidad en una forma confiable.

Flexibilidad significa ofrecer una amplia gama de productos y ser capaz de ajustarse a cambios inesperados en la demanda de la mezcla de productos que se ofrecen. Los fabricantes exitosos del siglo XXI serán aquellos que respondan más rápido a los cambios impredecibles en los gustos del cliente. El autor de este libro tuvo la fortuna de visitar la Planta Motomachi de Toyota en Ciudad Toyoda de Japón. Lo que le impresionó en particular fue la capacidad de producir diferentes modelos en la misma planta. De hecho, cada coche sucesivo en la línea de ensamble era de un modelo diferente. A un sedán Crown con el volante en el lado derecho, para el mercado doméstico, lo seguía un coupé Lexus con el volante del lado izquierdo, diseñado para embarque a Estados Unidos. Cada automóvil llevaba su propio conjunto de instrucciones que podían leer tanto robots de soldadura como ensambladores humanos. Esta flexibilidad permitió a Toyota ajustarse a la mezcla de producto en tiempo real, y lanzar un sistema en el que los clientes pueden ordenar automóviles con configuraciones personalizadas directamente desde terminales localizadas en las salas de exhibición de sus distribuidores (Port, 1999).

De hecho, una forma de concebir la estrategia de operaciones es el posicionamiento estratégico que la empresa selecciona para las dimensiones de costo, calidad, velocidad de entrega, confiabilidad de entrega y flexibilidad. La administración de las operaciones se relaciona con la implementación de la estrategia para lograr liderazgo a lo largo de estas dimensiones.

1.3 EL PUNTO DE VISTA CLÁSICO DE LA ESTRATEGIA DE OPERACIONES

La visión tradicional sobre la estrategia de manufactura fue desarrollada por Wickham Skinner de la Escuela de Negocios de Harvard, la cual ha sido mejorada posteriormente por varios investigadores, en su mayoría también de Harvard. El punto de vista tradicional trata la mayor parte de las cuestiones estratégicas en el contexto de una sola planta, más que en el de toda una compañía o corporativo. Los grandes temas comentados en la sección 1.2 se relacionan con la estrategia de operaciones a nivel de la compañía. Aunque puede considerarse que el punto de vista clásico está un poco pasado de moda, ciertas cuestiones siguen siendo importantes y relevantes. El pensamiento clásico de la estrategia de operaciones se relaciona con los temas siguientes:

1. Horizonte de tiempo
2. Focalización
3. Evaluación
4. Consistencia

Horizonte de tiempo

Horizonte de tiempo se refiere a la extensión de tiempo que se requiere para que la estrategia surta efecto. Una clasificación natural aquí consiste en separar las decisiones estratégicas en decisiones a corto, mediano y largo plazo. Las decisiones de operaciones a corto plazo pueden tener un impacto medible en días o incluso horas. Esto incluye decisiones concernientes a la compra, producción y planificación de personal; políticas para control de calidad y funciones de mantenimiento; cuestiones de control de inventario a corto plazo; programación de producción, inventario, y así sucesivamente. Las decisiones de corto plazo de este tipo conciernen al gerente de operaciones. Son el tema principal de este texto.

Las decisiones de operaciones a mediano plazo son aquellas que pueden medirse en términos de semanas y meses. Incluyen pronósticos de requerimientos y de demanda; decisiones de planeación de la mano de obra, incluyendo la determinación del tamaño y la mezcla de la fuerza de trabajo, y decisiones concernientes a la distribución de bienes a través de canales de distribución existentes y a establecer metas de la compañía en cuanto a niveles de inventario y de servicio. El cambio en los niveles que representa una señal de una situación fuera de control para una gráfica de control es una decisión relacionada con la calidad a corto plazo, mientras que reestructurar una línea de producción y buscar proveedores más confiables son cuestiones que pueden afectar la calidad del producto en un futuro posterior.

La estrategia se suele asociar con las decisiones de largo plazo. Elegir el momento oportuno en el tiempo, la ubicación y la escala de construcción de nuevas instalaciones de manufactura son decisiones típicas de la estrategia de operaciones y de manufactura a largo plazo. Tomar estas decisiones requiere de información sobre pronósticos de productos nuevos y existentes; sobre las cambiantes características del mercado, y sobre las modificaciones en los costos y la disponibilidad de recursos. La estrategia de manufactura debe constituir la base para la construcción de canales apropiados para las ventas y la distribución, así como para el desarrollo y diseño de instalaciones. La figura 1-1 muestra un desglose de las decisiones de operaciones y producción en términos de la distribución en el tiempo de las consecuencias que dichas decisiones tendrían.

Los horizontes de tiempo afectan el impacto de las decisiones, las incertidumbres que rodean esas decisiones y la penalización por haber tomado decisiones equivocadas. Los horizontes de corto plazo involucran muchas decisiones cuyo impacto puede ser pequeño pero que, al acumularse, pueden hacer la diferencia. Por ejemplo, los gerentes de las tiendas The Gap resurten todos los días sus anaqueles. El gerente hace pedidos por teléfono al centro de distribución (CD) y por lo general le llegan al día siguiente (suponiendo que el CD tenga en existencia los artículos solicitados). El gerente se basa en los datos de ventas y en el juicio del personal para realizar la selección de artículos que quiere reordenar. Los errores provocan faltantes en el inventario que resulta en ventas perdidas y molestia de los clientes, o sobrantes que no se venden, lo que resulta en un desperdicio de espacio en los anaqueles.

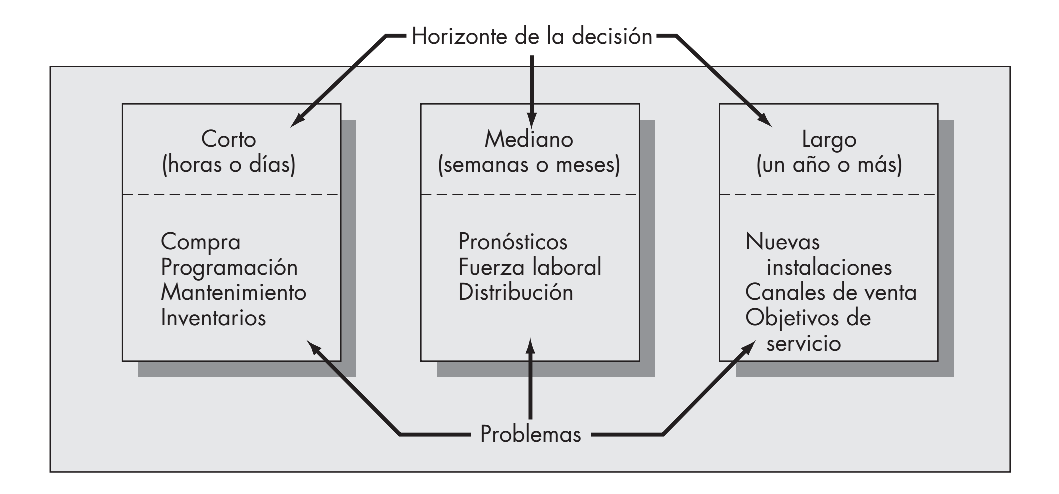


FIGURA 1-1
Horizontes de las decisiones de una estrategia de manufactura.

Los compradores de las oficinas generales de San Francisco deciden cuáles son las líneas de ropa que deben abastecer para la próxima temporada. Estas pueden considerarse decisiones de plazo intermedio. Los compradores poseen menos información para tomar sus decisiones que un gerente de una tienda, pero sus decisiones tienen mayor impacto. Los compradores deben juzgar las tendencias de la moda y las preferencias de color. Una línea que no se vende debe bajarse de precio y venderse con pérdida.

La alta dirección debe tomar decisiones de largo plazo. Para The Gap, algunas decisiones de largo plazo podrían ser: 1) el número, ubicación y tamaño de los centros de distribución; 2) los términos y condiciones para contratos a largo plazo con proveedores; 3) los arreglos para la logística de la cadena de abastecimiento para toda la compañía (interna *versus* externa), y 4) selección de personal directivo. En estas decisiones existe aún más incertidumbre. Un cambio demográfico puede dar como resultado que las decisiones tomadas previamente sobre la ubicación y el tamaño de los centros de distribución sea equivocada. Un contrato a largo plazo con una planta en el extranjero, por ejemplo, en China, puede resultar contraproducente. El gobierno estadounidense puede imponer algunas cuotas, o el gobierno chino puede imponer algún arancel. La importancia y sensibilidad de las decisiones a largo plazo probablemente son algunas de las razones por las que se paga tan bien al personal de la alta dirección en Estados Unidos.

Focalización

Skinner (1974) fue el primero en considerar la noción de la **focalización** o **enfoque** (del inglés ***focus***) **en la estrategia de manufactura**. Dio cinco características clave de la fábrica focalizada o enfocada (término conocido en inglés como *focused factory*):

1. *Tecnologías de proceso*. Una manera natural de realizar la focalización de las operaciones de una planta o fábrica se basa en el proceso que ésta emplea. La dirección debe limitar a una por fábrica aquellas tecnologías de proceso nuevas, que no hayan sido probadas, y mantener el número de las diferentes tecnologías de proceso maduras en un nivel en que el gerente de la fábrica pueda supervisarlas de manera eficiente.

2. *Demandas del mercado.* El mercado determina con frecuencia la focalización del producto o línea de productos que se generan en una fábrica. Las áreas típicas en las que el mercado dicta la focalización de una planta son:

- ***Precio.*** Hay evidencia de que el consumidor estadounidense está más consciente del precio que muchos consumidores del extranjero. El precio siempre ha sido un factor de diferenciación en Estados Unidos.
- ***Tiempo de demora.*** Los productos que no están protegidos por patentes deben producirse y distribuirse rápidamente para cubrir el mercado antes que sus competidores.
- ***Confiabilidad.*** Las especificaciones de confiabilidad difieren según el segmento de mercado, con frecuencia aun para productos idénticos. Como un ejemplo, consideremos una

compañía que está produciendo una línea de circuitos integrados que se venden a diferentes clientes. Uno lo usa en refrigeradores y otro en máquinas cardiopulmonares. Es posible que las especificaciones de confiabilidad deban ser mucho más estrictas en el caso de la máquina cardiopulmonar.

3. *Volúmenes de producto.* Los volúmenes de producción dentro de una misma planta deben resultar similares, de manera que el herramental, los sistemas de manejo de materiales y las líneas de producción de la planta no se subutilicen ni sobreutilicen.

4. *Nivel de calidad.* El nivel de calidad de los productos generados en una misma planta debe ser similar, a fin de que la compañía pueda establecer un estándar consistente de control de calidad. Los estándares de calidad son resultado de varios factores: las técnicas de control estadístico utilizadas, los procedimientos de monitoreo y las actitudes, procedimientos y capacitación de los trabajadores.

5. *Tareas de manufactura.* La productividad de una planta que fabrica una amplia línea de diferentes productos sufrirá interrupciones resultantes de las frecuentes preparaciones por cambio en los equipos y la reestructuración de sus líneas de producción. Cuando la dirección limita el número de las distintas tareas de manufactura por realizar en una misma localización, los trabajadores pueden concentrarse en perfeccionar los procesos existentes.

Existe evidencia de que las compañías focalizadas tienen mayor éxito. Hayes y Wheelwright (1984) relacionan el margen de utilidad operativa de 11 compañías con el número de tipos principales de productos que producen éstas. Con base en estos datos, muestran que las compañías que fabrican menos productos tienden a ser menos rentables. Es interesante subrayar que en las 11 compañías encuestadas la focalización, medida de esta manera, constituye un mejor predictor de la rentabilidad que el tamaño de las ventas.

Debe observarse que muchas compañías atienden diversos mercados de forma simultánea. Un ejemplo (y existen muchos más tanto en Estados Unidos como en Japón) es Yamaha. Yamaha se ha establecido como un fabricante de alta calidad de productos tan diversos como pueden ser equipos estereofónicos, instrumentos musicales (incluyendo instrumentos para bandas, pianos y sintetizadores digitales), así como equipo deportivo. Sin embargo, las fábricas focalizadas resultan preferibles aun para compañías que tienen una amplia línea de productos.

Evaluación

Existen varias dimensiones en las que se puede evaluar una estrategia de producción/operaciones. He aquí las más significativas:

1. *Costo.* En casos en los que el precio resulta clave para la competitividad y diferenciación de mercado, un medio importante para la evaluación de una estrategia es el costo de los productos que se entregan al cliente. Los costos directos de producción incluyen los costos de materiales, equipo y mano de obra. Los costos de distribución incluyen los costos asociados con el inventario, particularmente los de inventario en tránsito, de transportación y de distribución. Los costos administrativos de la planta y del proceso deben tenerse en cuenta también al calcular costos. Los costos administrativos de nuevos procesos pueden ser los más difíciles de evaluar debido a la incertidumbre tanto en la confiabilidad del proceso como en su vida útil.

2. *Calidad.* En los mercados donde la calidad del producto es un factor determinante para el éxito del producto en el mercado, o donde se requiere una alta confiabilidad para cumplir con las especificaciones del producto, es necesario evaluar una estrategia con base en la dimensión representada por la calidad. Por ejemplo, la calidad del producto es el medio principal para evaluar el desempeño de la manufactura en Japón.

3. *Rentabilidad.* Finalmente, es la rentabilidad de una línea de producto la que determina el éxito de una estrategia que se realiza con el fin de producirlo y venderlo. Sin embargo, como señalamos anteriormente, la maximización de utilidades a corto plazo puede ser una mala estrategia para la compañía si conlleva reducciones en la inversión en cuanto a nueva capacidad y tecnología. Si el horizonte de tiempo asociado con la evaluación de cualquier estrategia no es el correcto, tal vez la alta dirección estará tomando decisiones inadecuadas. Las estrategias que logran rentabilidad a corto plazo pueden no ser las mejores para los intereses a largo plazo de la compañía.

4. *Satisfacción del cliente.* Las compañías exitosas se han dado cuenta de que el éxito finalmente se obtiene sólo al lograr tener una base de clientes leales y satisfechos. Esto significa que el cliente no sólo debe estar satisfecho con un producto cuando lo compra, sino que debe confiar en que la compañía hará válidas las garantías proporcionando un servicio eficiente y efectivo en términos de costos después de realizada la venta. La *calidad regida por el mercado* es un concepto que escuchamos frecuentemente en nuestros días. Su reciente énfasis muestra que las compañías se están volviendo cada vez más conscientes de que el servicio al cliente debe ser una parte explícita del proceso de entrega del producto.

Consistencia

Con frecuencia se da el caso de que, más que tener una estrategia única consistente para la manufactura, la "estrategia" es meramente el conjunto agregado de todas las políticas de la compañía que impactan la manufactura. Las políticas salariales y de personal deben diseñarse para alentar la eficiencia y mejorar la productividad. El control de inventarios, la programación y los planes de producción se orientan a la minimización de costos y el mejoramiento de medidas como el tiempo ocioso del trabajador, el inventario de trabajo en proceso y los tiempos de demora de la producción. El diseño del proceso está dirigido a generar productos de alta calidad. El problema consiste en que cada una de estas metas individuales puede estar optimizando un objetivo diferente. El resultado es una compleja estructura de la planta en la que la dirección y la mano de obra asumen posiciones contrarias.

Skinner (1974) cita varias de las causas por las que se presentan inconsistencias comunes a la mayor parte de las compañías. Entre ellas podemos mencionar:

1. *Profesionalismo en la planta.* El número de diferentes cargos y funciones de trabajo se ha incrementado mientras que el alcance de éstos se ha reducido. Los profesionales de las diferentes áreas se interesan sólo en lucirse, y por lo tanto buscan maximizar sus contribuciones personales. Desafortunadamente, estos profesionales no necesariamente están trabajando para lograr las mismas metas. Algunos buscan minimizar costos, otros desean elevar la calidad del ambiente de trabajo y otros intentan optimizar el flujo de efectivo dentro de la compañía.

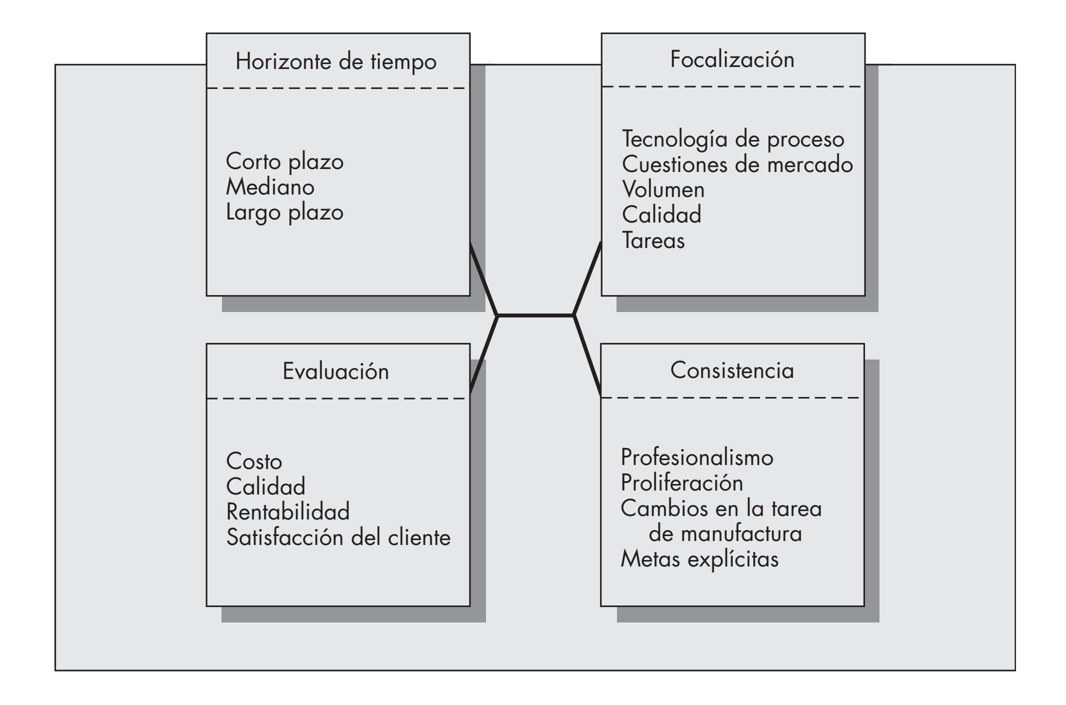
2. *Proliferación de productos.* Como se mencionó anteriormente, las compañías que producen un reducido número de productos tienden a ser más rentables. Conforme aumenta el número de productos diferentes, es más difícil conservar un conjunto de metas consistente dentro de la planta. ¿Por qué las compañías aumentan el tamaño de la línea de productos fabricada en una misma planta? Recordemos la discusión que tuvimos en la primera parte de este capítulo en torno a las consecuencias de medir el desempeño según el retorno de la inversión. Una forma segura de disminuir la inversión en nuevas capacidades es sencillamente producir nuevos productos en viejas plantas. Desde el punto de vista de la dirección, éste es un medio para reducir los gastos administrativos, pero también puede dar como resultado una baja calidad del producto y un control ineficiente de la producción.

3. *Cambios en la tarea de manufactura.* La dirección decide aumentar la capacidad productiva para poder responder a las necesidades dictadas por el mercado. Los mercados cambian, sin embargo, y como resultado de ello, pueden presentarse también cambios en el funcionamiento de la planta. Las metas y los medios para evaluar esas metas pueden no cambiar a la par que las demandas asignadas a la planta. Puede suceder que los objetivos que tenían sentido cuando se construyó la planta dejen de tenerlo tras cambiar el funcionamiento de la planta.

4. *La tarea de manufactura nunca se hizo explícita.* La dirección debe comprender la forma en que las metas de manufactura se alinean con la estrategia corporativa global para que surja una política consistente. Si no existe una estrategia corporativa clara, hay poca esperanza de que las áreas funcionales de la compañía —finanzas, mercadotecnia y producción— puedan desarrollar de forma independiente una estrategia consistente y significativa.

Los elementos esenciales de la discusión en esta sección se resumen en la figura 1-2.

FIGURA 1-2
Los elementos de una estrategia de producción y de operaciones.



1.4 COMPETIR EN EL MERCADO GLOBAL

La competencia internacional se ha convertido en una obsesión nacional. Los estadounidenses están preocupados porque su nivel de vida se está erosionando mientras en otros lugares del planeta parece mejorar. Existen evidencias de que hay algo de cierto en esta percepción. Nuestra balanza comercial con Japón ha estado en rojo durante décadas, y no hay ninguna señal de que las cosas vayan a cambiar. Las compañías estadounidenses que alguna vez tuvieron una posición de dominio a nivel mundial prácticamente han desaparecido al interior del país. Algunos ejemplos son los aparatos electrónicos para consumo doméstico, el acero y las herramientas eléctricas. Sin embargo, no todas las noticias son malas. La economía estadounidense es sólida y continúa creciendo. Las compañías estadounidenses siguen llevándose la tajada más grande del mercado mundial en muchas industrias.

En su excelente estudio sobre la competitividad internacional, Porter (1990) se pregunta: ¿por qué un determinado país se convierte en la base de operaciones de los competidores internacionales exitosos en una cierta industria? No puede discutirse el hecho de que ciertas industrias florecen en ciertos países. Algunos ejemplos son:

1. Alemania: máquinas impresoras, autos de lujo, productos químicos.
2. Suiza: productos farmacéuticos, chocolate.
3. Suecia: camiones pesados, equipo para minería.
4. Estados Unidos: computadoras personales, software, películas.
5. Japón: automóviles, artículos electrónicos para consumo doméstico, robótica.

¿Cómo se explica este fenómeno? Podemos ofrecer diferentes explicaciones, pero la mayoría tiene algo que las refuta. He aquí algunas:

1. *Historia.* Ciertas industrias son históricamente fuertes en algunos países y no se les puede desplazar con facilidad. *Refutación:* La desaparición de la industria siderúrgica en Estados Unidos refuta esta hipótesis.

2. *Estructura fiscal.* Algunos países, como Alemania, no imponen un impuesto a las ganancias de capital, proporcionando así un ambiente más fértil para la industria. *Refutación:* Sin embargo, no hay razón alguna para que un acuerdo fiscal benigno favorezca a ciertas industrias más que a otras.
3. *Características nacionales.* Muchos piensan que los trabajadores de otros países, particularmente de la cuenca del Pacífico, están mejor capacitados y son más dedicados que los trabajadores estadounidenses. *Refutación:* Si esto es cierto, ¿por qué entonces las compañías estadounidenses tienen mayor dominio en algunos segmentos de la industria? ¿Cómo podemos explicarnos el enorme éxito que han tenido las compañías japonesas que tienen plantas en Estados Unidos con fuerza laboral estadounidense?
4. *Recursos naturales.* No hay duda de que algunas industrias dependen en gran medida de los recursos naturales y de que estas industrias tienen una ventaja distintiva en algunos países. Un ejemplo es el de la industria de productos forestales en Estados Unidos y Canadá. *Refutación:* Muchos sectores de la industria son en esencia independientes de los recursos naturales, pero aun así florecen sólo en algunos países.
5. *Políticas gubernamentales.* Algunos gobiernos proporcionan asistencia directa a las industrias incipientes, como el MITI (Ministerio de la Industria y el Comercio Internacional, por sus siglas en inglés: *Ministry of International Trade and Industry*) en Japón. El papel del gobierno estadounidense es principalmente regulatorio. Por ejemplo, los estándares ambientales en los Estados Unidos probablemente son más estrictos que en cualquier otro lugar. *Refutación:* Esto no explica por qué algunas industrias dominan en países con estándares regulatorios y ambientales muy rígidos.
6. *Factores macroeconómicos ventajosos.* Tipo de cambio, tasas de interés y deuda gubernamental son algunos factores macroeconómicos que proporcionan una ventaja competitiva a los países. Por ejemplo, en los años 80, cuando las tasas de interés eran mucho más elevadas en Estados Unidos que en Japón, las compañías japonesas obtenían préstamos para nuevos proyectos con mayor facilidad. *Refutación:* Esos factores no explican por qué muchas naciones tienen un creciente nivel de vida pese a que sus déficit son cada vez mayores (Japón, Italia y Corea son algunos ejemplos).
7. *Mano de obra barata y abundante.* Aunque la mano de obra barata puede atraer nueva industria, la mayoría de los países con mano de obra barata son muy pobres. Por otro lado, muchos países (por ejemplo Alemania, Suiza y Suecia) tienen un elevado nivel de vida, altos salarios y escasez de mano de obra calificada.
8. *Prácticas gerenciales.* Hay evidencia de que las prácticas administrativas japonesas, en términos generales, son más efectivas que las occidentales. *Refutación:* Si las prácticas administrativas estadounidenses son tan poco efectivas, ¿por qué continúan dominando ciertas industrias, como las de computadoras personales, desarrollo de software y productos farmacéuticos?

Hablar de competitividad es más fácil que medirla. ¿Cuáles son las formas adecuadas para medir el éxito de un país en comparación con otro? Algunas posibilidades son

- Balanza comercial.
- Participación en las exportaciones mundiales.
- Creación de empleos.
- Bajos costos de mano de obra.

Pueden presentarse argumentos contra el uso de cada uno de los indicadores anteriores como una medida de la competitividad internacional. Suiza e Italia tienen déficit comercial y, al mismo tiempo, su nivel de vida ha crecido. Argumentos similares pueden hacerse en el caso de países que importan más de lo que exportan. El número de empleos creados gracias a la economía es un mal indicador de su nivel de bienestar. Es más importante la calidad de los empleos creados. Finalmente, los bajos costos de mano de obra se correlacionan con un bajo nivel de vida. Estas refutaciones muestran que no resulta fácil desarrollar una medida efectiva para la competitividad internacional.

Porter (1990) establece que la medida apropiada para comparar el desempeño nacional es la tasa de crecimiento de la productividad. Productividad es el valor de producto (*output*) por

unidad de insumo (*input*) de mano de obra o capital. Porter señala que el crecimiento de la productividad en algunas industrias parece ser más sólido en ciertos países, y que existen razones para ello. En algunos casos, podemos encontrar dichas razones en ciertas ventajas internas relacionadas con los diferentes factores. La teoría de los factores indica que todos los países tienen acceso a una misma tecnología (un supuesto que no es estrictamente verdadero) y que las ventajas nacionales surgen de los apoyos a los factores de producción como la tierra, la mano de obra, los recursos naturales y el capital.

He aquí algunos excelentes ejemplos de la teoría de los factores. Corea tiene costos de mano de obra relativamente bajos, así que exporta bienes que requieren de una intensiva mano de obra, como prendas de vestir y ensambles electrónicos. El mineral de hierro de Suecia es bajo en impurezas, lo que contribuye a la solidez de la industria siderúrgica sueca. Pese a ser tan persuasiva, la teoría de apoyo sobre los factores ha sido refutada. Por ejemplo, después de la guerra de Corea, Corea del Sur, aunque carecía de liquidez monetaria, desarrolló y se destacó en varias industrias altamente intensivas en el uso de capital, como el acero y la construcción naval. Además, muchos países tienen similars apoyos a los diferentes factores; sin embargo, algunos parecen sobresalir más en ciertas industrias. Estos ejemplos sugieren que los apoyos a los factores por sí solos no explican todos los casos de naciones con segmentos industriales dominantes.

Porter (1990) sugiere los siguientes cuatro elementos como determinantes de una ventaja nacional:

1. *Condiciones de factor* (ya comentadas).
2. *Condiciones de demanda.* Si los consumidores nacionales son sofisticados y exigentes, presionan a la industria local para que la misma innove con mucha rapidez, lo que da a las compañías una ventaja internacional. Los consumidores de artículos electrónicos en Japón son muy demandantes, lo que ha posicionado de manera muy competitiva a esta industria en el mercado internacional.
3. *Industrias relacionadas y de soporte.* Tener cerca a proveedores de clase mundial es una ventaja poderosa. Por ejemplo, la industria italiana del calzado se ve apoyada por una sólida industria de la piel y una bien establecida industria del diseño.
4. *Estrategia, estructura y rivalidad de la compañía.* La forma en que las compañías se organizan y administran contribuye a su competitividad internacional. El estilo administrativo japonés es marcadamente distinto del estadounidense. En Alemania, muchos altos ejecutivos cuentan con formación técnica, lo que genera una fuerte inclinación hacia el mejoramiento de productos y de procesos. En Italia existen muchas pequeñas compañías familiares, lo que alienta el individualismo.

Aun cuando Porter presenta un argumento muy convincente en cuanto a la ventaja competitiva de las naciones en algunas industrias, existe un debate entre los economistas acerca de si la noción de competitividad internacional tiene algún sentido. Son las compañías las que compiten, no los países. Tal es el punto de vista de Paul Krugman (1994). Según este último, Estados Unidos y Japón no son competidores de la misma forma en que lo son Ford y Toyota. El nivel de vida en un país depende de su propio desempeño económico interno y no de la forma en que se desempeña respecto a otros países.

Krugman sostiene que un énfasis excesivo en la competitividad internacional puede llevar a estrategias mal direccionadas. Este es muy probablemente el caso de las guerras comerciales. Así ocurrió cuando, a mediados de 1995, la administración de Clinton planeaba imponer altos aranceles a los fabricantes japoneses de autos de lujo. La mayoría de los economistas están de acuerdo con que las guerras comerciales y sus consecuencias, como los aranceles, a la larga no benefician a nadie. Otro problema que surge del orgullo competitivo nacional es que puede provocar gastos gubernamentales mal concebidos. Francia ha gastado miles de millones tratando de levantar su fallida industria de computadoras. (Ciertamente, no todas las inversiones gubernamentales en una industria doméstica pueden considerarse un error. El gobierno japonés, por ejemplo, jugó un importante papel en el apoyo a la industria de las pantallas planas. Las firmas japonesas dominan actualmente esta multimillonaria industria.)

Otro punto a favor de Krugman es que la mayor parte del producto interno bruto de Estados Unidos (PIB) se consume dentro de Estados Unidos, lo que hace que el éxito de una com-

pañía en nuestro mercado doméstico sea más importante que su éxito en el mercado internacional. Krugman admite que el crecimiento de la productividad es una preocupación válida. Argumenta, sin embargo, que debemos ser más productivos a fin de producir más, y no sólo por el hecho de ser mejores que nuestros competidores extranjeros.

El debate sobre la ventaja competitiva continuará. Los creadores de políticas deben estar conscientes de todos los puntos de vista, y sopesar con cuidado cada uno de ellos al formular una política. Aunque Krugman hace varios señalamientos interesantes, no hay duda de que la globalización es una tendencia que no muestra signos de revertirse. No podemos enterrar la cabeza en la arena y decir que los mercados internacionales carecen de importancia para nosotros. Las fronteras económicas están desapareciendo en todo el planeta.

Problemas para las secciones 1.1–1.4

1. ¿Por qué es indeseable que Estados Unidos evolucione hacia una economía de servicio?
2. ¿Cuáles desventajas ve usted si el CEO se preocupa principalmente por el rendimiento a corto plazo de una inversión?
3. ¿Recuerda alguna compañía que haya tenido que cerrar por enfocarse sólo en los costos sin lograr un estándar mínimo de calidad?
4. ¿Cuáles son los diferentes estándares de calidad a los que se hace referencia en el ejemplo que compara el Geo con el Ferrari?
5. Enumere los cuatro elementos de una estrategia de operaciones y comente el papel de cada uno en la planeación estratégica.
6. ¿Cuáles son las ventajas y desventajas de fabricar nuevos productos en instalaciones existentes?
7. Considere los siguientes escenarios industriales ficticios. En cada caso, identifique los posibles errores en la estrategia.

   a) Un fabricante de unidades (*drives*) para disco Winchester localizado en California está expandiendo su línea de productos para incluir discos ópticos. Ambos productos serán fabricados en las mismas instalaciones aun cuando los discos ópticos requieren de una tecnología de proceso diferente.

   b) En una fábrica de herramientas y troqueles en el Medio Oeste ha surgido un conflicto entre el jefe de control de calidad y el vicepresidente de operaciones. El gerente de control de calidad ha solicitado un nuevo equipo de fundición de moldes que tiene un costo de $400 000 y un excelente historial de confiabilidad, mientras que el vicepresidente de operaciones ha decidido comprar un producto similar que es menos confiable y el cual cuesta $300 000.

   c) El director de finanzas de la compañía ha establecido la política de no realizar nuevas inversiones en las instalaciones durante el siguiente año con el fin de mejorar el perfil actual de rentabilidad de la compañía.

   d) Un microprocesador diseñado en Estados Unidos y producido en una planta en México se utiliza para monitorear la presión de los fluidos. Se vende tanto a un fabricante de automóviles como a un fabricante de equipo para desarrollo sustentable.

8. ¿Cuáles son los cuatro determinantes de ventaja nacional que sugiere Porter? Proporcione ejemplos de compañías que han prosperado como resultado de cada uno de estos factores.
9. ¿Cuál ventaja de factor favorece a la industria del aluminio de Estados Unidos sobre Japón y hace que resulte mucho más barato producirlo aquí? (Sugerencia: La producción de aluminio requiere de mucha energía. ¿En qué parte del país se dispone de una fuente de energía barata?)
10. Paul Krugman sostiene que como la mayor parte del producto interno se consume nacionalmente, no deberíamos hacer demasiado hincapié en la competencia internacional. ¿Qué

Aplicación panorámica

READ-RITE GANA CON UNA ALIANZA ESTRATÉGICA INTERNACIONAL

Read-Rite Corporation, fundada en 1983 y con sede en Milpitas, California, fue uno de los fabricantes independientes estadounidenses más grandes de cabezas para unidades (*drives*) de disco. La cabeza es un circuito integrado que hace posible la lectura y escritura de datos en los medios magnéticos. La tecnología de las cabezas es compleja y volátil; la cabeza debe rediseñarse con cada nueva generación de unidades de disco.

Construir nuevas instalaciones de manufactura resulta extremadamente costoso. Durante los ochenta, la compañía luchó por conservar el paso en la cambiante tecnología de los discos duros. Para los noventa la compañía fue rentable de nuevo, pero muchos problemas continuaron. Uno de ellos era que tenían un solo cliente (Periféricos Conner). Otros dos problemas con los que se toparon fueron disponer de liquidez para la expansión y conservar la delantera ante el rápido desarrollo de la competencia japonesa.

Para enfrentar el avance de los japoneses, era importante competir con ellos en su propio terreno, pensó el CEO Cyril Yansouni. Una fuerte demanda interna (como comentó Porter) es un determinante crucial para la ventaja competitiva nacional. El acceso a los mercados japoneses siempre es difícil para la gente de fuera, así que Read-Rite buscó una alianza estratégica con un socio japonés.

El socio que eligió Read-Rite fue Industrias Metálicas Sumitomo. Aparentemente esta fue una selección un poco extraña. Sin embargo, aunque Sumitomo no tenía presencia en tecnología en ese momento, la compañía estaba llevando a cabo varias inversiones en negocios basados en tecnología. También, como la tercera empresa siderúrgica de Japón, Sumitomo disponía de mucho efectivo. Como parte del acuerdo, Sumitomo convino en pagar $30 millones a cambio de una participación accionaria en Read-Rite.

El mayor riesgo que enfrentó Read-Rite fue la pérdida de su tecnología ante un competidor potencial. Para protegerse, de acuerdo con Yansouni, "íbamos a hacer que un eventual divorcio fuera muy doloroso para ambas partes". Ninguna tendría derecho a vender cabezas en Japón sino hasta muchos años después de la separación, momento en el cual cualquier tecnología adquirida seguramente sería para entonces obsoleta.

¿Cómo resultaron las cosas? Con la infusión de efectivo proveniente de Sumitomo, Read-Rite pudo atraer nuevos clientes en Estados Unidos, incluyendo a Western Digital, Quantum y Maxtor. El año siguiente, la compañía se volvió pública, con efectivo extra proveniente de la oferta pública inicial; tuvo suficiente dinero para comenzar a construir plantas ensambladoras en Asia. La alianza (conocida como Read-Rite SMI) tuvo mucho éxito en Japón y logró alcanzar el segundo puesto en ventas de cabezas en Japón, sólo detrás de TDK. El éxito de la alianza estratégica ha ayudado a suavizar algunas de las fluctuaciones en los ingresos nacionales de Read-Rite y ha constituido una parte muy importante del éxito de la compañía en los años noventa. Casi una década después, la alianza estratégica continúa triunfando en Japón, a pesar de los problemas económicos que allá se registran.

Fuente: "Pacific Overtures", *The Institutional Investor*, enero de 1996.

Nota del autor: Aunque esta alianza estratégica probó ser valiosa para ambas partes, Read-Rite se declaró en quiebra en 2003. Sus activos fueron adquiridos por Western Digital Corporation.

industrias en Estados Unidos han sido más golpeadas por la competencia extranjera? ¿Cuáles son las amenazas potenciales para Estados Unidos si el conjunto de estas industrias fracasa?

11. Krugman señala algunos programas gubernamentales mal orientados que han surgido derivados de un excesivo énfasis en la competitividad internacional. ¿Qué riesgos pueden aparecer si se pone poco énfasis en la competitividad internacional?

1.5 INICIATIVAS ESTRATÉGICAS: REINGENIERÍA DE PROCESOS DE NEGOCIOS

Como si estuvieran programados, cada cierto número de años surge un nuevo método de control de producción o una técnica de administración, casi siempre descrito con siglas de tres letras. Aunque es fácil ser escéptico, en general los métodos son buenos y pueden tener un valor sustancial para las compañías cuando se implementan de manera inteligente. La *reingeniería de procesos de negocios* (BPR, por sus siglas en inglés: *Business process reengineering*) surgió tras

la publicación del libro de Hammer y Champy (1993). La BPR no es una técnica específica, como puede ser la planeación de requerimiento de materiales, ni un concepto de planeación de la producción como el justo a tiempo (*just-in-time*). En vez de ello, se centra en la idea de que es posible cambiar y mejorar los procesos de negocios establecidos. Se trata de cuestionar por qué las cosas se hacen de una manera, sin aceptar como respuesta "porque ésa es la forma en la que lo hacemos".

Hammer y Champy, que definen la BPR como "comenzar de nuevo" proporcionan varios ejemplos de esfuerzos exitosos de reingeniería. El primero es IBM Credit Corporation (la Corporación de Crédito de IBM), una compañía subsidiaria de IBM que, si fuera independiente, aparecería en la lista de las 100 más grandes compañías de servicio de *Fortune*. Esta división de IBM es responsable de proporcionar crédito a nuevos clientes que adquieren equipo IBM. El proceso tradicional de aprobación de crédito seguía cinco pasos:

1. Un vendedor de IBM pide telefónicamente una solicitud de financiamiento; la solicitud se registra en una hoja de papel.
2. Alguien sube el documento al departamento de crédito, donde otra persona ingresa la información en una computadora y revisa la calificación crediticia del solicitante del préstamo. El especialista escribe los resultados de la revisión de crédito en un papel, que se envía al departamento de prácticas comerciales.
3. Un tercero, en el departamento de prácticas comerciales, actualiza el formato estándar del crédito en respuesta a la solicitud del cliente. Estas actualizaciones, que se realizan en otro sistema de cómputo, se anexan a la solicitud original y al informe del especialista del departamento de crédito.
4. La solicitud pasa a un fijador de precios, que introduce la información en una hoja de cálculo para determinar la tasa de interés que se cargará al cliente. La recomendación del fijador de precios se escribe en papel y se envía (junto con el resto de los documentos) al personal administrativo.
5. La información se convierte en una cotización que será entregada al vendedor por medio de Federal Express.

Este proceso requería un promedio de seis días y a veces hasta dos semanas. Los representantes de ventas registraban un sinnúmero de quejas por este retraso: durante este tiempo, el cliente puede encontrar otra fuente de financiamiento u otro vendedor. En un esfuerzo para probar si este proceso podía simplificarse, dos directivos de alto nivel decidieron hacer el seguimiento de una nueva solicitud a lo largo de los cinco pasos, pidiéndole al personal que dejara a un lado lo que estaba haciendo y lo procesara como normalmente lo haría. Descubrieron que para llevar a cabo el proceso completo de cinco pasos se requería de ¡únicamente 90 minutos en promedio! El resto del tiempo, o bien las solicitudes estaban en tránsito de un departamento a otro, o esperaban en el escritorio de alguna persona para ser procesadas. Claramente, el problema no estaba en la eficiencia del personal sino en el diseño del proceso mismo de autorización de crédito.

La solución resultó sencilla: los cuatro especialistas que manejaban cada solicitud de crédito fueron reemplazados por un generalista en créditos que manejaba la solicitud de principio a fin. Se diseñó un software de actualización para apoyar al generalista, quien no tuvo ningún problema en manejar la mayoría de las solicitudes. El proceso de aprobación de crédito original se había diseñado asumiendo que cada solicitud era lo suficientemente compleja para requerir de alguien con cierto conocimiento en cada una de las áreas. En realidad, la mayoría de las solicitudes eran cuestión de rutina, y los especialistas generalmente no tenían más que consultar una sencilla tabla para determinar la cifra adecuada.

¿Cuál fue el resultado de este cambio? ¡Los seis días que se requerían para responder las solicitudes de crédito se convirtieron en sólo cuatro horas! Y esto se logró con menos personal y con un incremento de cien veces más contratos manejados.

Aunque cada esfuerzo de reingeniería debe pensarse con cuidado y no hay dos soluciones exactamente iguales, Hammer y Champy (1993) sugieren que los esfuerzos de reingeniería utilicen los principios generales que se presentan a continuación:

1. *Varios trabajos se combinan en uno solo.* Pocos ejemplos de BPR son tan impresionantes como el de IBM Credit, pero hay otras historias de éxito en la literatura. Muchos casos de éxito tienen una característica común: la reducción de un proceso complejo que requiere muchos pasos para convertirlo en uno más simple que requiere de menos pasos. En el caso de IBM Credit, un proceso de cinco pasos se redujo a uno solo. Esto sugiere un principio general. El proceso de IBM Credit fue una evolución natural del concepto de la división del trabajo. El economista Adam Smith expuso este principio en el siglo XVIII (véase la cita de *La riqueza de las naciones* al principio de la sección 1.10 de este capítulo). Sin embargo, algo que puede ser bueno puede llevarse demasiado lejos. Si un proceso se divide en demasiados pasos, finalmente se llega al punto de rendimientos decrecientes. Los éxitos más importantes que alcanzó la BPR provienen de simplificar procesos complejos reduciendo el número de pasos requeridos.

2. *Los trabajadores toman decisiones.* Nuestra meta es reducir el número de niveles a los que se debe de reportar, permitiendo a los trabajadores tomar decisiones que antes estaban reservadas a la dirección. En el caso de IBM Credit, la mayoría de las decisiones alguna vez reservadas a los especialistas ahora las toma un solo generalista. Dar a los trabajadores mayor poder de decisión puede suponer una amenaza para la dirección, que puede ver ese paso como una intromisión en sus prerrogativas.

3. *Los pasos en el proceso se realizan siguiendo un orden natural.* Los pasos de un proceso no necesariamente deben efectuarse siguiendo una secuencia lineal rígida, sino en un orden tal que tenga sentido en el contexto del problema que se está resolviendo. En particular, en muchos casos, algunas tareas pueden realizarse de forma simultánea más que en secuencia. (Estas ideas, desde luego, son bien conocidas y forman la base para los conceptos de la administración de proyectos del capítulo 8.)

4. *Los procesos deben tener múltiples versiones.* Se deben considerar diversas contingencias, no con base en el diseño de múltiples procesos independientes, sino diseñando un proceso lo suficientemente flexible, capaz de reaccionar ante diferentes circunstancias. En el caso de IBM Credit, por ejemplo, el proceso de emisión de crédito final tenía tres versiones: una para los casos directos (manejados por computadora), una para los casos de dificultad media (manejados por el estructurador del trato) y otra para los casos difíciles (realizado por el estructurador del contrato con la ayuda de los asesores especialistas).

5. *El trabajo se realiza donde tiene más sentido hacerlo.* Uno de los principios básicos de la reingeniería es no llevar demasiado lejos la idea de la división del trabajo. Otro es no llevar demasiado lejos la idea de la centralización. Por ejemplo, en la mayoría de las compañías las compras se realizan de manera centralizada. Esto significa que cada solicitud de compra está sujeta a los mismos esfuerzos administrativos en cuanto a tiempo y papeleo. ¡Una consecuencia puede ser que el costo de procesar una solicitud exceda el costo del artículo que se va a adquirir! En el caso de artículos de bajo costo, puede ahorrarse mucho si se permite a cada departamento manejar sus propias compras. (Hammer y Champy tratan un caso como ese.)

Los autores enumeran otros principios básicos, que involucran la minimización de verificaciones y conciliaciones; el tener un punto de contacto único, y el ser capaces de emplear operaciones híbridas en cuanto a la descentralización/centralización.

Es más fácil enumerar los pasos que uno debe considerar en un esfuerzo de reingeniería, que implementarlo en la realidad. En el mundo real, no es posible ignorar la realidad política. Para muchas de las historias de éxito en la literatura, no sólo se simplificaron los procesos, sino que también se tuvo una reducción del personal. Resulta entendible que los empleados vean la BPR como una excusa velada para recortar personal (a la que, con un eufemismo, se denomina “ajuste de personal”). Este fue exactamente el caso de una compañía de servicios financieros. Cuando se filtró que la dirección estaba planeando un esfuerzo de reingeniería, la mayor parte del personal asumió que ocurrirían despidos masivos. Algunos incluso pensaron que la compañía estaba camino a la quiebra. Por su parte, los líderes sindicales vieron la reingeniería como un medio de la alta dirección con el que se buscaba desaparecer las categorías laborales y las

reglas de trabajo que habían ganado tras duras negociaciones a lo largo de años, y persuadieron a los miembros del sindicato para iniciar una huelga. En un tercer caso, un gerente general estaba molesto con la potencial pérdida de autoridad que podía acompañar un esfuerzo de reingeniería. Renunció para abrir su propio negocio. (Estos ejemplos se presentan en un libro de seguimiento escrito por Hammer y Stanton, 1995.)

Tales historias muestran que iniciar un esfuerzo de reingeniería no carece de riesgos. En muy pocas ocasiones el proceso resulta tan sencillo como en IBM. La reingeniería ha sido descrita por Ronald Compton, el CEO de Aetna Life y Casualty, como "angustiante, desgarrador, duro". Debe realizarse un análisis de costo-beneficio para estar seguros de que las ganancias potenciales compensan los riesgos.

La optimización de procesos no es nueva. En sus primeros años, el campo de la ingeniería industrial se encargó del diseño óptimo de los procesos, estableciendo estándares a partir del uso de estudios de tiempos y movimientos, y con el uso de diagramas de flujo para comprender la secuencia de los acontecimientos y el flujo de los materiales en una fábrica. ¿Cómo difiere de la BPR? En primer lugar, la BPR aborda los flujos de los procesos de negocio más que de los flujos del proceso de manufactura. En segundo lugar, el concepto no se refiere a la optimización de un proceso existente, sino a repensar, partiendo de cero, la forma en que deben hacerse las cosas. Como tal, es más revolucionaria que evolutiva. Tal vez sea muy perturbadora, pero puede traer consigo grandes beneficios. Para que funcione la BPR, los empleados de todos los niveles tienen que hacer suyo el enfoque, y la alta dirección debe estar al frente. De otra manera, el esfuerzo de reingeniería puede resultar en un costoso fracaso.

1.6 INICIATIVAS ESTRATÉGICAS: JUSTO A TIEMPO (JUST-IN-TIME)

El *justo a tiempo* (*JIT*, por sus siglas en inglés) es un proceso de manufactura, por un lado, y una estrategia de operaciones amplia, por el otro. Los elementos de JIT se comentarán con mayor detalle en el capítulo 7, como parte de un análisis completo de los sistemas de inventarios de empujar y jalar. Sin embargo, JIT (o producción esbelta, como también se le conoce) es una filosofía que incluye el tratamiento de inventarios en la planta, las relaciones con los proveedores y las estrategias de distribución. El núcleo de la filosofía es eliminar desperdicios. Esto se logra mediante una programación eficiente de los pedidos que entran, los inventarios de trabajo en proceso y los inventarios de bienes terminados.

JIT se deriva del **sistema kanban** introducido por Toyota. Kanban es una palabra japonesa que significa tarjeta o boleta. Originalmente, las tarjetas kanban fueron sólo un medio para implementar el JIT. El sistema kanban fue introducido por Toyota para reducir el exceso en inventarios de trabajo en proceso (*WIP*, por sus siglas en inglés: work-in-process). Hoy en día, JIT es mucho más ambicioso. Tanto los sistemas de control de calidad como las relaciones con los proveedores son parte de un sistema JIT integrado. Los sistemas JIT pueden implementarse en formas distintas a las que utilizan las tarjetas kanban. Al integrar las filosofías JIT con sistemas de información sofisticados se logra que la transferencia de información sea mucho más rápida. La velocidad en la que puede transferirse la información de una parte de la compañía a otra es un factor importante para el éxito del sistema JIT.

JIT es una filosofía para operar una compañía que incluye el establecimiento de relaciones de entendimiento y de trabajo con proveedores, previendo un cuidadoso monitoreo de la calidad y el flujo de trabajo, y asegurando que los productos sean fabricados únicamente conforme se van necesitando. Aunque JIT puede usarse meramente en la forma como la diseñó originalmente Toyota, es decir, como un medio para mover el trabajo en proceso de un centro de trabajo a otro, los defensores del método recomiendan mucho más: integrar la filosofía JIT en toda su estrategia de negocio.

Los sistemas de flujo de material e inventarios se clasifican en **sistemas de empujar** (también conocidos por su formulación en inglés: *push*) o **sistemas de jalar** (conocidos también por el término común en inglés: *pull*). Un sistema de empuje es aquel en el que las decisiones relativas a la forma en que fluirá el material a través del sistema se toman de manera centralizada.

Con base en estas decisiones, el material se produce y "empuja" el siguiente nivel del sistema. Un sistema típico de empuje es la planeación de requerimientos de material (MRP, por sus siglas en inglés) que se comentará con detalle en el capítulo 7. En MRP, las cantidades apropiadas de producción para todos los niveles de la jerarquía de producción se determinan todas al mismo tiempo, con base en los pronósticos de demanda del producto final y la relación entre los componentes y los artículos finales. En JIT, la producción se inicia en un determinado nivel a consecuencia de una solicitud realizada por un nivel superior. Las unidades, entonces, se "jalan" a través del sistema.

JIT tiene muchas ventajas sobre los sistemas convencionales. La eliminación de los inventarios WIP reduce costos de mantener inventarios. Un menor inventario significa menos dinero atado al inventario. JIT también permite detectar con rapidez problemas en la calidad. Ya que las unidades se fabrican sólo conforme se van necesitando, en un sistema JIT que esté funcionando adecuadamente nunca debe darse la situación en la que se generan grandes cantidades de inventario WIP defectuoso antes de que se pueda detectar la existencia de un problema de calidad. JIT igualmente implica estrechar las relaciones con los proveedores. Los proveedores deben estar dispuestos a absorber algo de incertidumbre y ajustar las cantidades y tiempos de entrega para acoplarse con las tasas de flujo de los productos.

Parte de lo que hizo que el sistema kanban fuera tan efectivo para Toyota fue su éxito en la reducción de tiempos de preparación de las operaciones críticas. El ejemplo más impresionante de reducción del tiempo de preparación es el llamado SMED (por sus siglas en inglés: *single-minute exchange of dies*), o intercambio de dados en cuestión de un minuto. Cada vez que se inicia un cambio importante en el estilo del cuerpo de una parte, es necesario cambiar los dados utilizados en el proceso. La operación típica de cambio de dados solía llevar de cuatro a seis horas. Era necesario parar toda la línea de producción durante la operación de cambio de dados. La dirección de Toyota se enteró de que Mercedes Benz fue capaz de reducir su operación de cambio de dados a menos de una hora. Al darse cuenta de que era posible realizar reducciones todavía más espectaculares, Toyota se enfocó en la reducción de tiempo requerido para el cambio de dados. En una serie de impresionantes mejoras, Toyota finalmente redujo esta operación crítica a sólo unos cuantos minutos. La idea esencial detrás de SMED consiste en realizar el mayor número posible de cambios fuera de línea (*off-line*), mientras continúa el proceso de producción.

Un aspecto importante de JIT consiste en el establecimiento de relaciones con los proveedores. ¿Qué distingue a las compras JIT de las prácticas de compras convencionales? Freeland (1991) da una lista de las características contrastantes entre la conducta JIT y la convencional. Algunas de éstas incluyen:

Compra convencional	**Compra JIT**
1. Entregas grandes, poco frecuentes	1. Entregas pequeñas con mucha frecuencia
2. Múltiples proveedores para cada parte	2. Pocos proveedores; una sola fuente
3. Contratos de compra a corto plazo	3. Contratos a largo plazo
4. Mínimo intercambio de información	4. Intercambio frecuente de información
5. Precios determinados por los proveedores	5. Precios negociados
6. Proximidad geográfica poco relevante	6. Proximidad geográfica relevante

En su estudio, Freeland observa que las industrias que parecían beneficiarse más de las compras JIT eran las que tienen inventarios muy grandes. Las compañías sin compras JIT se orientaban a una configuración tipo taller o a la fabricación sobre pedido. Los proveedores que realizaban acuerdos de compras JIT tendieron a incrementar sus inventarios de seguridad, lo cual sugiere que los fabricantes están reduciendo sus inventarios a costa de los proveedores. La separación geográfica entre proveedores y compradores representaba un serio obstáculo para implementar con éxito las compras JIT. La industria automotriz fue una de las que reportó un mayor número de beneficios provenientes de los convenios de compras JIT. En otras industrias, como la de las computadoras, se tuvieron respuestas mezcladas; algunas compañías reportaron beneficios sustanciales, mientras que otras reportaron muy pocos beneficios.

Aunque la reducción del exceso de inventario de trabajo en proceso puede traer muchos beneficios, JIT no necesariamente es la respuesta a todas las situaciones de manufactura. De acuerdo con Stasey y McNair (1990),

> el inventario en una planta típica es como un seguro, garantiza que el problema en un área de la planta no va a afectar el trabajo realizado por otra. Cuando se solucionan los problemas que crean la necesidad de un seguro, los inventarios desaparecen de la planta.

Si eliminamos todas las fuentes de incertidumbre en la planta, desaparece la necesidad de tener inventarios. El problema es que algunas fuentes de variación nunca pueden eliminarse. Una sería la demanda del consumidor. JIT es efectivo sólo si la demanda final es regular. Otras pueden ser las fuentes de variación inherentes al proceso de producción o del equipo. ¿Se puede simplemente establecer por decreto la desaparición de todas las fuentes de incertidumbre en el ambiente de manufactura? Desde luego que no. De hecho, aunque los principios subyacentes del JIT son muy sólidos, no es la cura para todos los males y no necesariamente será el método correcto para todas las situaciones de producción.

1.7 INICIATIVAS ESTRATÉGICAS: COMPETENCIA BASADA EN EL TIEMPO

El profesor Terry Hill de la Escuela de Negocios de Londres (*London Business School*) ha propuesto una interesante forma de ver los factores competitivos. Los clasifica en dos tipos: "calificadores" y "ganadores de pedidos". Un producto que no posea un factor de calificación debe dejar de tomarse en cuenta. El ganador de pedidos es el factor que determina quién se queda con las ventas de entre el campo de los calificadores.

Dos factores de los que siempre se hace mención son la calidad y el **tiempo hacia el mercado** (conocido en inglés como *time to market*). En la década pasada, los japoneses y alemanes se ganaron la lealtad de muchos consumidores estadounidenses gracias a la calidad de sus productos. Las compañías estadounidenses están poniéndose al día en la dimensión de calidad. A partir de lo que se estableció en la sección 1.6, vemos que las compañías estadounidenses exitosas han sido capaces de producir productos que igualan las tasas de defectos de los competidores extranjeros. Si continúa esta tendencia, el consumidor dará por hecho la calidad del producto. La calidad se convertirá entonces en un factor calificador, más que en uno ganador de pedidos.

Si es éste el caso, ¿cuáles son los factores que determinarán a los ganadores de pedidos en los próximos años? Las compañías automotrices japonesas fabricaron y continúan fabricando automóviles de alta calidad. En años recientes, sin embargo, los principales productores de automóviles en Japón han comenzado a enfocarse en la estética y los gustos del consumidor. Se han diversificado y han pasado del poco emocionante mercado de los autos compactos de los setenta y ochenta a nuevos mercados como la línea de lujo Lexus de Toyota, y el innovador y exitoso Miata de Mazda.

La introducción oportuna de características y diseño novedosos determinará a los ganadores de pedidos en la industria automotriz. En cuanto a las computadoras, Compaq construyó su reputación en buena medida a partir de su capacidad para ser la primera en llevar nueva tecnología al mercado. La competencia basada en el tiempo es un término del que escucharemos cada vez con más frecuencia en los próximos años.

¿Qué significa **competencia basada en tiempo**? No se trata de los estudios de tiempos y movimientos populares en los años treinta que conformaron la base de la disciplina de ingeniería industrial. En vez de eso, según Blackburn (1991),

> Los competidores basados en el tiempo se enfocan en lograr una amplia perspectiva, en todo el sistema de generación y entrega de valor. Buscan transformar toda una organización en una entidad focalizada en el tiempo total que se requiere para entregar un producto o servicio. Su meta no es encontrar la mejor forma de realizar una tarea, sino eliminar ésta completamente o realizarla en paralelo con otras tareas, de manera que se reduzca el tiempo total de respuesta del sistema en su conjunto. Para convertirse en un competidor basado en el tiempo se requiere de cambios revolucionarios en las formas en que se organiza el proceso.

Los minoristas exitosos entienden muy bien la competencia basada en el tiempo. El éxito de las cadenas de ropa de moda The Gap y The Limited se debe en gran medida a su capacidad de brindar al cliente lo último en moda de una manera oportuna. Parte del éxito de la cadena sumamente exitosa de Wal-Mart es su estrategia de administración del tiempo. Cada artículo en las existencias de Wal-Mart se reabastece dos veces por semana, mientras que el promedio para la industria es de una vez cada dos semanas. Esto hace posible que Wal-Mart logre mejores tasas de rotación de inventarios que su competencia, y le permite responder con mayor rapidez a las demandas de los clientes. Las estrategias de Wal-Mart la han llevado a convertirse en líder de la industria, con una tasa de crecimiento tres veces más grande que el promedio de la industria, y lograr utilidades dos veces por encima de sus competidores (Blackburn, 1991, capítulo 3).

La administración basada en el tiempo es un elemento de mucha complejidad para los manufactureros, y en algunas industrias resulta obvio que es un factor clave para el éxito o el fracaso. Los líderes de la industria de la memoria de acceso aleatorio dinámico (DRAM, por sus siglas en inglés: *dynamic random access memory*) cambiaron cuatro veces entre 1978 y 1987. En cada caso, la compañía que fue la primera en llevar al mercado la siguiente generación de DRAM fue la que dominó el mercado. La experiencia DRAM se resume en la siguiente tabla (Davis, 1989):

Producto	Compañía	Año de introducción	Primer año de producción masiva	Líderes del mercado en el primer año de producción masiva
16K	Mostek	1976	1978	Mostek (25%) NEC (20%)
64K	Hitachi	1979	1982	Hitachi (19%) NEC (15%)
256K	NEC	1982	1984	NEC (27%) Hitachi (24%)
1 MB	Toshiba	1985	1987	Toshiba (47%) Mitsubishi (16%)

Desconozco algún otro ejemplo que muestre con tanta claridad y tan predeciblemente el valor de ser el primero en entrar al mercado.

1.8 INICIATIVAS ESTRATÉGICAS: COMPETENCIA EN CALIDAD

¿Cuáles son los factores competitivos que los directivos estadounidenses consideran que serán relevantes durante la próxima década? Basados en una encuesta realizada a 217 participantes de la industria, los siguientes factores se consideraron como los más importantes para lograr una ventaja competitiva en los próximos años. Están enumerados en orden de importancia.

1. Calidad de la conformancia
2. Desempeño en la oportunidad de la entrega
3. Calidad
4. Flexibilidad del producto
5. Servicio posventa
6. Precio
7. Variedad de una línea (características)
8. Distribución
9. Flexibilidad del volumen
10. Promoción

En esta lista vemos algunos temas importantes. La **calidad** y la **administración del tiempo** aparecen como factores líder. El control de calidad ha recibido mucha difusión en los últimos

años tras el establecimiento del prestigioso premio Malcolm Baldridge (creado basándose en el premio Deming japonés, que tiene mucho más tiempo). Calidad significa diferentes cosas en situaciones distintas, por lo que apremia darle sentido según el contexto de los bienes manufacturados. Un producto de alta calidad es aquel que funciona de acuerdo con la manera en que fue diseñado. Los productos funcionarán según la forma en la que fueron diseñados para funcionar si se presenta sólo una ligera variación en su proceso de manufactura. Esta definición permite que un producto con un diseño pobre pueda resultar con una alta calidad, mientras que también puede resultar posible que un producto con un muy buen diseño pueda resultar de baja calidad. Aun aceptando esta estrecha definición de calidad, ¿cuál es la mejor medida de la misma? Las tasas de defecto son un barómetro típico. Sin embargo, una medida más adecuada puede ser la confiabilidad del producto después de su manufactura. Esta medida se usa normalmente para monitorear la calidad de productos como automóviles y artículos electrónicos.

Durante los años recientes, en Estados Unidos se ha despertado un enorme interés en torno a la calidad. Con la invasión de la competencia japonesa, muchas industrias estadounidenses están luchando por sobrevivir. El negocio de vender calidad está en el punto más elevado de su historia. Las consultoras especializadas en programas de calidad para la industria, como el Juran Institute y Philip Crosby Associates, están teniendo un enorme éxito. La cuestión es si las compañías estadounidenses sólo están hablando de dientes para afuera respecto a la calidad, o si en verdad tratan de modificar la forma de hacer negocios. Existe evidencia de que, cuando menos en algunos casos, esto último es la verdad.

Por ejemplo, si comparamos las compañías de automóviles estadounidenses y japonesas, en las estadounidenses la calidad medida por los defectos reportados por los propietarios de autos durante sus primeros tres meses disminuyó significativamente entre 1987 y 1990, lo que redujo notablemente la brecha con Japón. La División Buick de General Motors, un ganador del premio Baldridge, ha realizado increíbles mejoras en este sentido. Entre 1987 y 1990, Buick disminuyó su tasa de defectos en 70%, igualando el porcentaje de Honda en 1990 (*Business Week*, 22 de octubre de 1990).

Existen muchas historias de éxito en la industria manufacturera norteamericana. Ford Motors logró triunfar con el Taurus. Mejoró tanto la calidad como la innovación, proporcionando a los compradores autos confiables y de tecnología avanzada. En 1980, James Harbour reportó que los fabricantes de autos japoneses eran capaces de producir un coche a un costo de $1 500 inferior que sus contrapartes estadounidenses. Ford ha cerrado esta brecha a unos cuantos cientos de dólares. Parte del éxito de Ford radica en la decisión de su antiguo CEO Donald Peterson, que consiste en no invertir millardos de dólares en nuevas plantas que incorporasen la tecnología más novedosa como lo hizo GM a mediados de los ochenta. Pero esto es sólo parte de la historia. De acuerdo con Faye Wills (1990),

> Si desea buscar respuestas sorpresivas para el ascenso de Ford, si se buscan sus secretos ocultos, hay que olvidarse de eso. El truco consiste en una buena y sólida administración del día con día: ejecución y planeación "siguiendo el librito", aplicar el sentido común a los procedimientos y los diseños de planta, diseños inteligentes que además de vender autos, permitan reducir costos e incrementen los márgenes de utilidades. Es así de sencillo.

Podemos aprender de nuestros propios éxitos. La industria de herramientas fue una de aquellas en las que más incursionaron los japoneses en los ochenta. Muchas compañías estadounidenses se vieron derrotadas por el ataque de la competencia asiática, pero no Stanley Works de New Britain, Connecticut. En 1982, el presiente de la compañía estaba analizando si Stanley debía permanecer en el negocio de artículos de ferretería tras la avalancha de las compañías asiáticas en el mercado estadounidense con martillos, desarmadores y otras herramientas de bajo precio. Stanley decidió contraatacar. Modernizó sus plantas e introdujo nuevos sistemas de control de calidad. Entre 1982 y 1988 las tasas de desperdicio (*scrap*) bajaron de 15 a sólo 3% en New Britain; Stanley no sólo enfrentó a la competencia, sino que también compitió exitosamente en Asia. Stanley ahora tiene una muy rentable operación vendiendo sus muy características cintas métricas amarillas en Asia.

¿Dónde se fabrica el mayor número de clones de PC? ¿Taiwán? ¿Corea? Es necesario considerar otra opción. La respuesta puede resultar toda una sorpresa: Texas. Dos firmas tejanas

han resultado sumamente exitosas en este mercado. Una es Computadoras Compaq (ahora parte de Hewlett Packard), que ingresó en el mercado a principios de los ochenta con la primera computadora personal portátil. Continuó elaborando productos con muy buenos diseños y alta calidad, y logró dominar 20% del mercado mundial de computadoras personales. Compaq se consolidó como un líder en el mercado. Otro fabricante exitoso de computadoras personales, Dell Computer, también es de Texas. El súbito crecimiento de la firma Dell resulta interesante. Michael Dell, un ex alumno de la Universidad de Texas, comenzó revendiendo computadoras personales IBM a principios de los ochenta. Más tarde creó PC'S Limited, que comercializó uno de los primeros clones de computadoras personales por correo. Dell es un líder del mercado en el área de las computadoras personales; ofrece una combinación de diseños vanguardistas, productos de alta calidad y excelente servicio.

Otra firma estadounidense que se ha comprometido seriamente con la calidad es Motorola. Motorola, ganador del premio Baldridge en 1987, ha reducido consistentemente la tasa de defectos en los productos que manufactura. Los defectos reportados alcanzaban casi 40 partes por millón a finales de 1991; bajaron de 6 000 partes por millón reportadas en 1986. Motorola anunció que su meta era alcanzar el 6-sigma (lo que significa seis desviaciones estándar respecto a la media de una distribución normal), lo que se traduce a 3.4 partes por millón. Motorola siente que el proceso de buscar el premio Baldridge fue tan valioso que ahora exige a sus proveedores competir por el mismo.

Historias de éxito como estas muestran que Estados Unidos pueden competir de manera exitosa con Japón y otros rivales extranjeros en cuanto a la calidad. No obstante, si realmente queremos tener clase mundial, la administración de la calidad total debe arraigarse en nuestra cultura. Se debe contar con sistemas que permitan monitorear las medidas tradicionales de calidad: ajustarse a las especificaciones y productos sin defectos. Sin embargo, la administración de la calidad debe ir más allá de las medidas estadísticas. La calidad requiere compenetrarse en la forma en que hacemos negocios, desde la calidad en diseño, calidad en manufactura y calidad para crear sistemas de trabajo con los proveedores, hasta la calidad en cuanto a satisfacción y servicio al cliente.

Problemas para las secciones 1.5 a 1.8

12. ¿Cuál es una definición operacional de calidad? ¿Es posible que una televisión de 13 pulgadas que se vende en $100 tenga mejor calidad que una de 35 pulgadas que cuesta $1 800?
13. Diferentes estudios han mostrado que las tasas de defectos de muchos productos japoneses son mucho menores que las de sus contrapartes estadounidenses. Especule sobre las posibles razones de estas diferencias.
14. ¿Qué significa "competencia basada en el tiempo"? Dé un ejemplo de un producto que haya sido introducido al mercado antes que sus competidores.
15. Considere el refrán "El que pega primero pega dos veces". Comente el significado de esta frase en el contexto de la competencia basada en el tiempo. En particular, ¿ser el primero en el mercado representa el único factor a considerar para el éxito de un producto?
16. ¿Cuáles características generales buscaría usted en un proceso de negocios que permitirían considerarlo como un candidato para la reingeniería? Comente una situación de su propia experiencia en la que resulte claro que un determinado proceso de negocios podría haberse mejorado.
17. ¿En qué formas pueden ser útiles las siguientes técnicas para un esfuerzo de reingeniería?
    - Simulación en computadora
    - Diagramado de flujos
    - Técnicas de la administración de proyectos
    - Modelado matemático
    - Equipos transfuncionales

18. ¿Qué problemas puede prever que se presentarán en las siguientes situaciones?
    a) La alta dirección está interesada en la reingeniería para lograr reducir costos, pero los empleados muestran escepticismo.
    b) Los trabajadores de línea desearían que se implementara un esfuerzo de reingeniería, porque lo consideran un medio para ser tomados más en consideración en la toma de decisiones, pero la dirección no muestra ningún interés.
19. El justo a tiempo ha sido caracterizado como un sistema cuya meta primordial es eliminar el desperdicio. Comente de qué manera se puede presentar desperdicio en *a*) las relaciones con los proveedores, *b*) la recepción de materia prima en la planta y *c*) el movimiento de materiales a través de la planta. ¿Cómo reducen los métodos JIT estas diferentes formas de desperdicio?
20. ¿En qué forma pueden mejorar los sistemas JIT la calidad de un producto?

1.9 ACOPLAMIENTO DE LOS CICLOS DE VIDA DEL PRODUCTO Y DEL PROCESO

El ciclo de vida del producto

La demanda de nuevos productos normalmente experimenta ciclos que pueden identificarse y mapearse a lo largo del tiempo. Comprender la naturaleza de esta evolución ayuda a identificar las estrategias apropiadas para la producción y las operaciones en las diversas etapas del ciclo del producto. Un **ciclo de vida del producto** típico se muestra en la figura 1-3. El ciclo de vida del producto estaría formado por cuatro segmentos principales:

1. Inicio
2. Rápido crecimiento
3. Maduración
4. Estabilización o declive

Durante la fase de inicio, el mercado para el producto se está desarrollando, los costos de producción y distribución son elevados y, usualmente, la competencia no es un problema. Durante

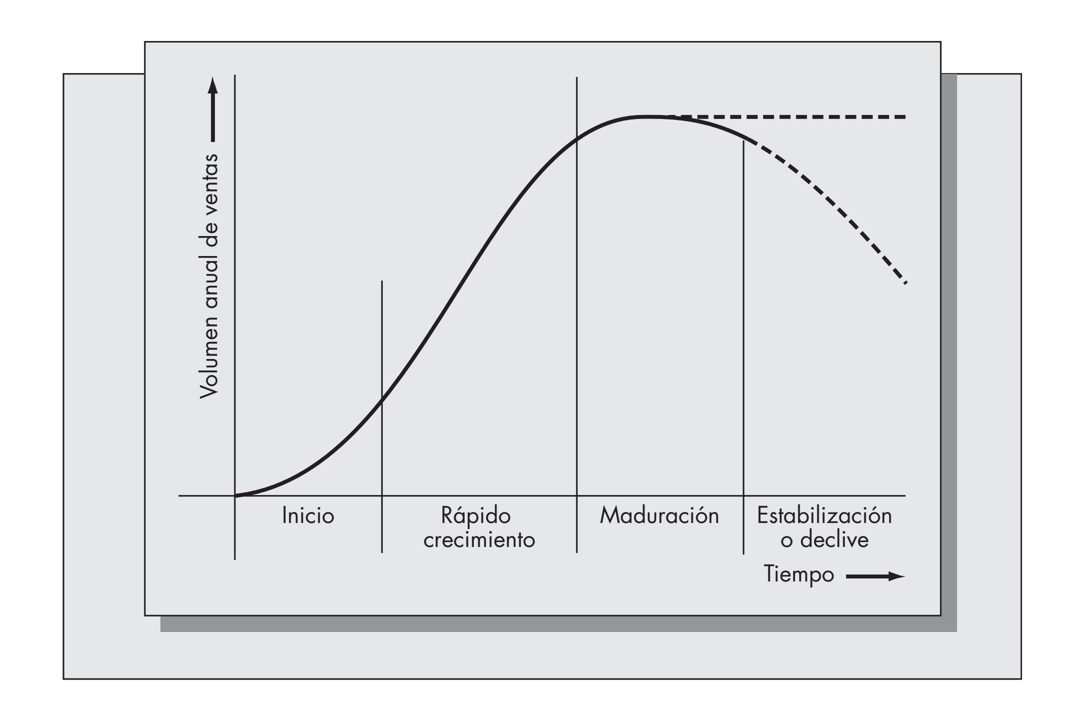


FIGURA 1-3
La curva del ciclo de vida del producto.

esta fase la estrategia primaria consiste en ocuparse de aplicar las experiencias del mercado y de manufactura para mejorar las funciones de producción y mercadotecnia. En este momento, deben manifestarse y corregirse las deficiencias serias de diseño.

El periodo de crecimiento rápido atestigua el principio de la competencia. La meta estratégica primaria durante este periodo consiste en posicionar en el mercado el producto tan firmemente como sea posible. Para lograrlo, la dirección debe considerar patrones alternativos de asignación de precios que se adapten a las diferentes clases de clientes y que refuercen la preferencia de la marca entre proveedores y consumidores. El proceso de manufactura debe llevar a cabo mejoras y acciones de estandarización conforme aumentan los volúmenes de producción. La flexibilización y modularización de la función de manufactura es altamente deseable en esta etapa.

Durante la fase de maduración del ciclo de vida del producto, el objetivo debe ser mantener y mejorar la lealtad a la marca que la compañía ha ido cultivando durante la fase de crecimiento. La dirección debe buscar incrementar la participación de mercado estableciendo precios competitivos. Los ahorros en los costos deben realizarse a través de un mejor control de la producción y la distribución del producto. Durante esta fase la compañía necesita atender las señales que envíe el mercado. La mayoría de los problemas con la calidad y el diseño del producto deben corregirse durante las fases de inicio y crecimiento, pero también deben contemplarse mejoras adicionales durante esta fase.

La forma apropiada de la curva del ciclo de vida para la etapa final depende de la naturaleza del producto. Muchos productos continuarán vendiéndose con un potencial anual de crecimiento casi de forma indefinida. Ejemplos de dichos productos pueden ser artículos electrodomésticos, alimentos procesados y automóviles. En lo tocante a estos productos, las principales metas de la compañía en esta fase, en esencia serán iguales a las descritas previamente para la fase de maduración. Otros productos experimentarán un declive natural en el volumen de ventas conforme se satura el mercado del producto o conforme el producto se va volviendo obsoleto. Si este es el caso, la compañía debe adoptar una estrategia para sacar el mayor provecho posible del producto o la línea de producto al mismo tiempo que minimiza la inversión en nueva tecnología de manufactura y en publicidad en medios masivos de comunicación.

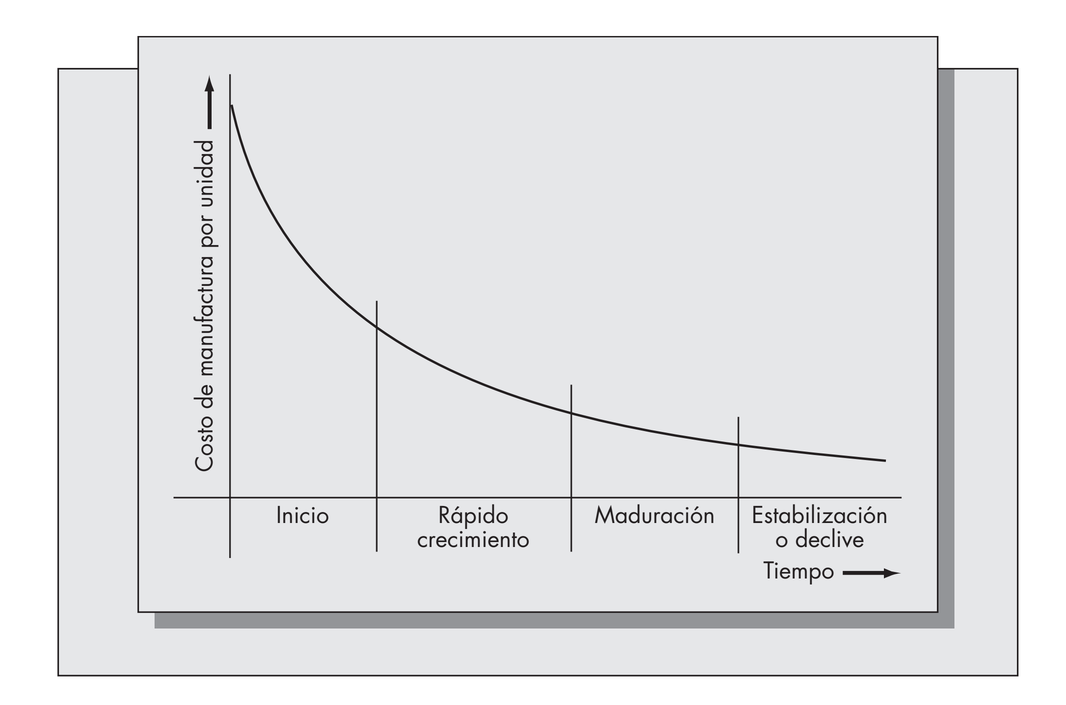
Aunque es un concepto útil, la curva del ciclo de vida del producto no es del todo preciso para todas las circunstancias. Los departamentos de mercadotecnia que basan sus estrategias en la curva del ciclo de vida pueden tomar decisiones erróneas. Dhalla y Yuspeh (1976) mencionan el caso de una compañía que cambió el peso de su inversión en publicidad de un producto exitoso estable a un nuevo producto. Supusieron que el nuevo producto estaba entrando en la fase de crecimiento de su ciclo de vida y el producto estable iniciaba la fase de declive de su ciclo de vida. Sin embargo, los consumidores nunca aceptaron el nuevo producto y, debido a la caída en el presupuesto asignado para su publicidad, las ventas del producto estable declinaron y nunca se recuperaron. Ello sugiere que, en algunas circunstancias, resulta más práctico adoptar un modelo consistente con la historia del producto y con el comportamiento del consumidor, más que asumir ciegamente que todos los productos deben seguir el mismo patrón de crecimiento y declive. Aun cuando consideramos que, en general, el concepto del ciclo de vida es una forma útil de observar los patrones de demanda de los consumidores, la construcción cuidadosa de un modelo para cada producto representa una muy efectiva herramienta de planificación.

El ciclo de vida del proceso

Abernathy y Townsend (1975) han clasificado en tres las etapas principales del **ciclo de vida del proceso de manufactura**: temprana, media y madura. Estas fases no necesariamente coinciden de forma precisa con las etapas del ciclo de vida del producto, pero proporcionan un marco conceptual para planear mejoras en el proceso de manufactura conforme madura el producto.

En la primera fase del ciclo de vida del proceso, la función de manufactura tiene las características de un taller (*job shop*). Debe enfrentar una mezcla diversificada de pedidos de relativamente bajo volumen y ser capaz de responder a cambios en el diseño del producto. Los tipos y calidad de los insumos pueden variar considerablemente, y la compañía tiene poco control sobre los proveedores.

FIGURA 1-4
El ciclo de vida del proceso y la curva de la experiencia.



En la fase media del ciclo de vida del proceso, la automatización comienza a jugar un papel más importante. La compañía debe ser capaz de ejercer mayor control sobre los proveedores conforme se incrementa el volumen de producción. Los costos unitarios de producción disminuyen como resultado de los efectos de aprendizaje. El proceso de producción puede involucrar el procesamiento por lotes (*batch*) y algunas líneas de transferencia (líneas de ensamble).

En la última fase del ciclo de vida del proceso, se automatizan la mayoría de las operaciones principales, se estandariza el proceso de producción y se introducen muy pocas innovaciones a la manufactura. El proceso de producción puede asumir las características de una operación de flujo continuo.

Este escenario evolutivo no es adecuado para todo nuevo proyecto de manufactura. Por ejemplo, una compañía que florece en un contexto de pedidos pequeños y con características peculiares deberá conservar las características de una organización tipo taller. El concepto de ciclo de vida del proceso se aplica sólo a nuevos productos que finalmente tenderán a madurar en artículos con alto volumen de producción. El problema de acoplar las características del producto con las características del proceso se aborda más adelante.

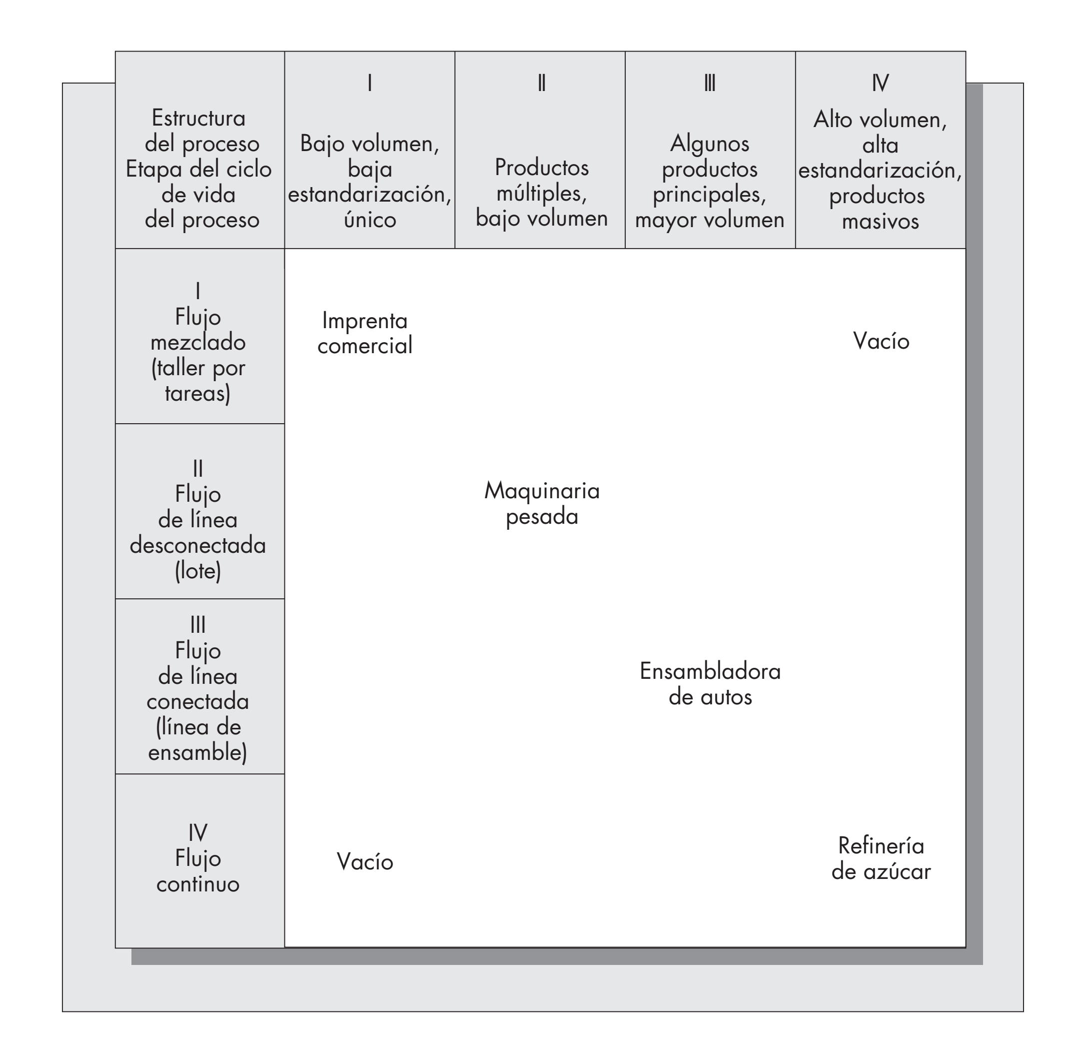
Las curvas de la experiencia muestran que los costos unitarios de producción disminuyen conforme se incrementa el número acumulado de unidades producidas. Se puede visualizar la curva de la experiencia en términos del ciclo de vida del proceso mostrado en la Figura 1-4. Una visualización precisa de la relación entre la curva de la experiencia y el ciclo del proceso de vida puede resultar invaluable. Al acoplar la disminución en el costo unitario con las diferentes etapas del ciclo de vida del proceso, la dirección puede compenetrarse mejor en las consecuencias de pasar de una fase del ciclo de vida del proceso a otra. Ello contribuirá a determinar los tiempos más adecuados para llevar a cabo mejoras en el proceso de manufactura.

La matriz producto-proceso

Hayes y Wheelwright (1979) consideran la relación entre los ciclos de vida del producto y del proceso mediante la **matriz de producto-proceso** ilustrada en la figura 1-5. La matriz se basa en cuatro fases en la evolución del proceso de manufactura: 1) flujo mezclado, 2) flujo de línea desconectada, 3) flujo de línea conectada y 4) flujo continuo. Esta matriz puede verse de dos maneras. Una consiste en acoplar la industria apropiada en su fase de madurez con el correspondiente proceso apropiado. Este punto de vista reconoce que no todas las industrias necesa-

FIGURA 1-5
La matriz de producto-proceso.

Fuente: Robert H. Hayes y Steven C. Wheelwright, "Link Manufacturing Process and Product Life Cycles", en *Harvard Business Review*, enero-febrero de 1979. © 1979 por el presidente y miembros colegiados del Harvard College; derechos reservados. Reimpreso con autorización.



riamente siguen la evolución del proceso descrita en la sección previa al ciclo de vida del proceso. Ciertas compañías o ciertos productos pueden permanecer indefinidamente en una etapa del ciclo de vida del proceso. Sin embargo, aun las compañías que no evolucionan a una posición en la esquina inferior derecha de la matriz deben, en la mayoría de los casos, ser ubicadas en alguna parte de la diagonal de la matriz.

Localizadas en la esquina superior izquierda de la matriz se encuentran las compañías que se especializan en trabajos únicos en su tipo, en los que la función de manufactura tiene las características de un almacén de flujo mezclado. Una imprenta comercial ejemplifica un almacén de flujo mezclado. La producción se realiza en lotes relativamente pequeños, y la organización es de tipo taller para obtener el máximo de flexibilidad.

Hacia abajo de la diagonal principal se encuentran las compañías que requieren una gran flexibilidad para producir una línea reducida de artículos estandarizados. Los fabricantes de maquinaria pesada caen en esta categoría debido a que deben producir en volúmenes algo elevados. Una línea desconectada proporcionará la suficiente flexibilidad para cumplir con los pedidos personalizados al mismo tiempo que conserva las economías de una estandarización reducida.

La tercera categoría hacia abajo de la diagonal incluye compañías que producen una línea de productos estándar para un mercado de grandes volúmenes. Los ejemplos típicos son productores de electrodomésticos o equipo electrónico, y fabricantes de automóviles. La línea de ensamble o línea de transferencia será una tecnología de proceso adecuada en este caso.

Finalmente, la porción inferior derecha de la matriz será adecuada para productos que involucran un flujo continuo. El procesamiento de productos químicos, gasolina y refinamiento de petróleo y refinamiento del azúcar son algunos ejemplos. Dichos procesos se caracterizan por bajos costos unitarios, estandarización del producto, altos volúmenes de ventas e inflexibilidad extrema del proceso de producción.

¿Cuál es el punto medular de este esquema de clasificación específico? Proporciona un medio para evaluar si la compañía opera en la porción adecuada de la matriz; esto es, si el proceso está adecuadamente acoplado con la estructura del producto. Las compañías que eligen operar fuera de la diagonal deben comprender claramente las razones para hacerlo. Un ejemplo de una empresa exitosa que opera fuera de la diagonal sería Rolls Royce. Otra es una compañía que fabrica muebles hechos a mano. El proceso de manufactura en estos casos deberá tener las características del almacén de flujo mezclado, pero sus competidores pueden ubicarse normalmente en la segunda o tercera posición de la diagonal.

Existe otra forma de ver la matriz producto-proceso. Puede usarse para identificar el acoplamiento adecuado entre el proceso de producción y las fases del ciclo de vida del producto. En la fase de inicio del desarrollo del producto, la compañía deberá posicionarse normalmente en la esquina superior izquierda de la matriz. Conforme madura el mercado para el producto, la compañía se moverá hacia abajo a lo largo de la diagonal para lograr economías de escala. Finalmente, la compañía se quedará en la posición de la matriz que sea adecuada tomando como base las características específicas del producto.

Problemas para la sección 1.9

21. *a*) ¿Cuáles son las cuatro fases del proceso de manufactura que aparecen en la matriz producto-proceso?

    b) Comente las desventajas de operar fuera de la diagonal de la matriz.

22. Proporcione un ejemplo de un producto que haya pasado por las cuatro fases del ciclo de vida del proceso y haya logrado estabilidad.
23. Discuta la siguiente aseveración: "Todas las compañías deben evolucionar a lo largo de la diagonal de la matriz producto-proceso".
24. Ubique las siguientes operaciones en la posición adecuada de la matriz producto-proceso.

    a) Un pequeño taller de reparación de instrumentos musicales.
    b) Una refinería de petróleo.
    c) Un fabricante de mobiliario para oficina.
    d) Un fabricante de línea blanca para el hogar como lavadoras, secadoras y refrigeradores.
    e) Una compañía manufacturera en la fase de inicio.

1.10 CURVAS DE LA EXPERIENCIA Y EL APRENDIZAJE

Conforme se adquiere experiencia en la producción de un producto específico, sea a nivel de un trabajador o por una industria en su conjunto, el proceso de producción se va tornando más eficiente. Como observó el economista Adam Smith en el siglo XVIII en su histórica obra *La riqueza de las naciones*:

> La división del trabajo, mediante la reducción de las labores de una persona a una única operación, y hacer de esta operación su única ocupación en la vida, necesariamente incrementa en gran medida la destreza del trabajador.

Al cuantificar la relación que describe la ganancia en eficiencia que se obtiene con el incremento en el número acumulado de unidades producidas, la dirección puede predecir de manera precisa la capacidad de las instalaciones existentes y los costos unitarios de producción. Hoy en día reconocemos que hay muchos otros factores, además del mejoramiento de las habilidades de una persona, que contribuyen a lograr este efecto. Algunos de estos factores incluyen:

- Mejoramiento de los métodos de producción.
- Mejoramiento en la confiabilidad y eficiencia de las herramientas y máquinas utilizadas.
- Mejor diseño del producto.
- Mejor programación de la producción y el control de los inventarios.
- Mejor organización del lugar de trabajo.

Diversos estudios de la industria aeronáutica realizados durante los años veinte mostraron que las horas de mano de obra directa que se requerían para producir una unidad disminuían con el incremento en el número acumulado de unidades producidas. El término **curva de aprendizaje** se adoptó para explicar este fenómeno. Igualmente se ha observado que en muchas industrias los costos marginales de producción disminuyen con el incremento en el número acumulado de unidades producidas. El término **curva de la experiencia** se ha usado para describir este segundo fenómeno.

Curvas de aprendizaje

Conforme los trabajadores adquieren más experiencia con los requerimientos de un proceso particular, o conforme el proceso se va mejorando con el tiempo, disminuye el número de horas necesario para producir una unidad adicional. La curva de aprendizaje, que modela esta relación, es también un medio para describir economías de escala dinámicas. La experiencia ha mostrado que estas curvas se pueden representar de manera precisa por medio de una relación exponencial. Sea $Y(u)$ el número de horas de mano de obra que se requiere para producir la u-ésima unidad. Entonces la curva de aprendizaje es

$$Y(u) = au^{-b},$$

donde a es el número de horas que se requieren para producir la primera unidad y b mide la tasa marginal de disminución de las horas de producción con el incremento en el número de unidades producidas. Tradicionalmente, las curvas de aprendizaje se describen por medio de la baja porcentual de las horas de mano de obra requeridas para producir el artículo número $2n$ comparado con las horas de mano de obra necesarias para producir el artículo número n, y se asume que este porcentaje es independiente de n. Esto es, una curva de aprendizaje de 80% significa que el tiempo que se requiere para producir la unidad número $2n$ es 80% del tiempo requerido para producir la unidad número n para todo valor de n. Para una curva de aprendizaje de 80%

$$\frac{Y(2u)}{Y(u)} = \frac{a(2u)^{-b}}{au^{-b}} = 2^{-b} = .80.$$

Se desprende que

$$-b \ln(2) = \ln(.8)$$

$$\text{o } b = -\ln(.8)/\ln(2) = .3219. \text{ (ln es el logaritmo natural).}$$

En términos más generales, si la curva de aprendizaje es una curva de aprendizaje de $100L$ por ciento, entonces

$$b = -\ln(L)/\ln(2).$$

La figura 1-6 muestra una curva de aprendizaje de 80%. Cuando se grafica en un papel con doble logaritmo o log-log, la curva de aprendizaje debe ser una línea recta si la relación exponencial que hemos asumido es precisa. Si se toman los logaritmos de ambas partes de la expresión $Y(u)$, se obtiene como resultado una relación lineal, ya que

$$\ln(Y(u)) = \ln(a) - b \ln(u).$$

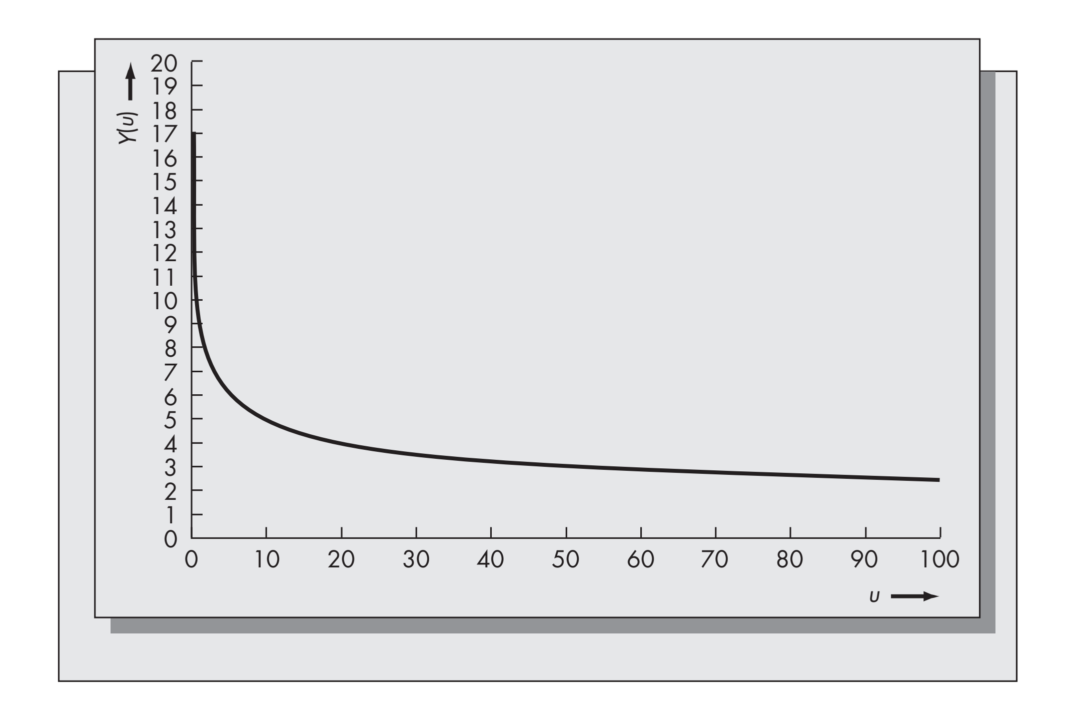


FIGURA 1-6
Una curva de aprendizaje de 80%.

Se utiliza regresión lineal para ajustar los valores de a y b a los datos reales una vez hecha la transformación logarítmica. (En el apéndice 2-B se presentan ecuaciones generales para encontrar los estimadores de mínimos cuadrados en regresión lineal.)

Ejemplo *1.1*

XYZ ha conservado registros cuidadosos del número promedio de horas de mano de obra que se requieren para producir uno de sus nuevos productos, un transductor de presión utilizado en los sistemas de combustible del automóvil. Estos registros aparecen en la siguiente tabla.

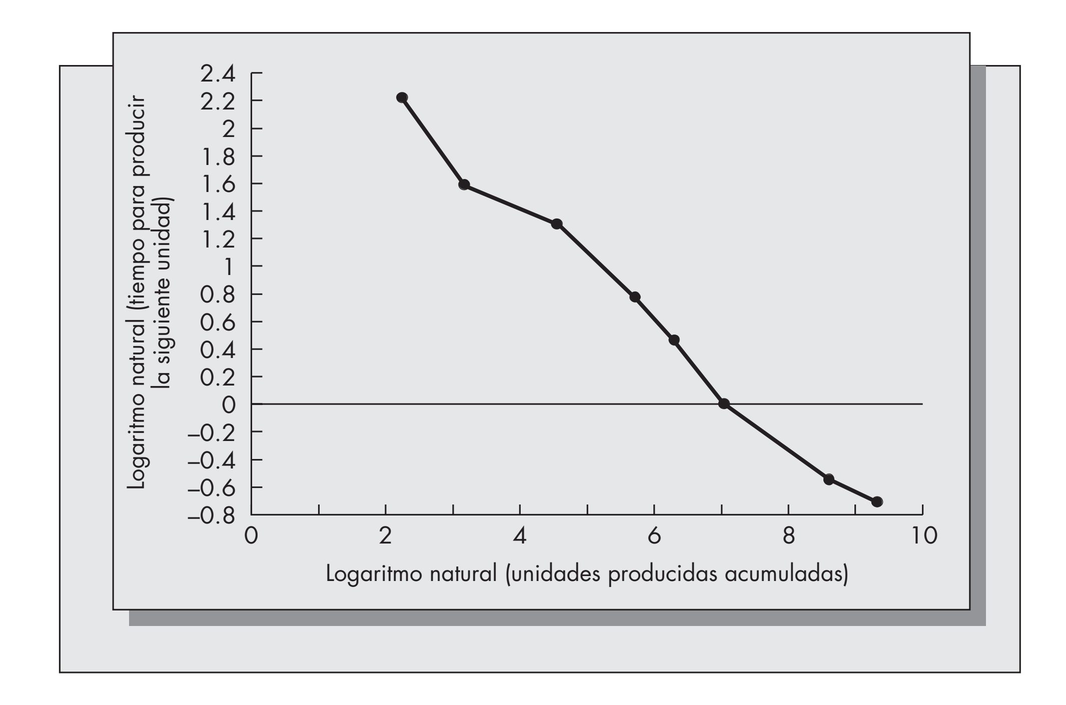
De acuerdo con la teoría, debe existir una relación lineal entre el logaritmo del número acumulado de unidades producidas y el de las horas requeridas para la producción de la última unidad producida. La gráfica de los logaritmos de estas cantidades para los datos anteriores se muestra en la figura 1-7. La figura sugiere que la curva exponencial de aprendizaje es bastante precisa para este caso. Utilizando los métodos delineados en el apéndice 2-B, hemos obtenido estimadores para la pendiente y la ordenada al origen del ajuste de mínimos cuadrados para los datos en la figura 1-7. Los valores de los estimadores de mínimos cuadrados son:

$$\text{Ordenada al origen} = 3.1301$$
$$\text{Pendiente} = .42276.$$

Número acumulado de unidades producidas (A)	ln (Columna A)	Horas requeridas para la siguiente unidad (B)	ln (Columna B)
10.00	2.30	9.22	2.22
25.00	3.22	4.85	1.58
100.00	4.61	3.80	1.34
250.00	5.52	2.44	0.89
500.00	6.21	1.70	0.53
1 000.00	6.91	1.03	0.53
5 000.00	8.52	0.60	−0.51
10 000.00	9.21	0.50	−0.69

Debido a que la ordenada al origen es ln(a), el valor de a es exp(3.1301) = 22.88. En consecuencia, la primera unidad debe haber requerido 23 horas para ser producida. El término para la pendiente es la

FIGURA 1.7
Gráfica log-log de los datos de XYZ.



constante $-b$. De la ecuación para b en la página anterior tenemos que

$$\ln(L) = -b\ln(2) = (-.42276)(.6931) = -.293.$$

Se desprende que $L = \exp(-.293) = .746$.

Así, estos datos muestran que el efecto del aprendizaje sobre la producción de los transductores puede describirse en forma precisa mediante una curva de aprendizaje de 75%. Esta curva se puede usar para predecir el número de horas de mano de obra que se requerirán para la posterior producción de estos transductores. Por ejemplo, al sustituir u = 50 000 en la relación

$$Y(u) = 22.88u^{-.42276}$$

se tiene un valor de Y(50 000) = .236 horas. Sin embargo, dichos resultados deben interpretarse con cautela. Una relación de la curva de aprendizaje puede no ser válida indefinidamente. Eventualmente, el producto puede alcanzar el final de su ciclo de vida natural, lo que podría ocurrir antes de que se hayan producido las 50 000 unidades de este ejemplo. Por otro lado, puede haber un límite absoluto para el número de horas de mano de obra que se necesitan para producir una unidad que, debido a la naturaleza del proceso de manufactura, eventualmente ya no puede mejorarse. Aun teniendo en mente estas limitaciones, las curvas de aprendizaje pueden ser una valiosa herramienta de planeación cuando se utilizan de manera adecuada.

Curvas de la experiencia

Las curvas de aprendizaje pueden ser un medio para calibrar la reducción de las horas de mano de obra marginales conforme los trabajadores se familiarizan con una tarea en particular o con el incremento en la eficiencia en el proceso de producción. Las curvas de la experiencia miden el efecto que la experiencia acumulada con la producción de un producto o familia de productos tiene sobre los costos globales y sobre el precio. Las curvas de la experiencia resultan más valiosas para industrias que están pasando por cambios sustanciales, como son la industria de la microelectrónica, más que para industrias muy maduras en las que ya han ocurrido los cambios más radicales, como la industria automotriz. La baja constante en los precios de los circuitos integrados es un ejemplo clásico de una curva de la experiencia. La figura 1-8 (Noyce, 1977) muestra el precio unitario promedio como una función de la experiencia acumulada de la industria medida en millones de unidades producidas durante el periodo de 1964 a 1972. Esta gráfica se muestra en una escala log-log y los puntos están muy cercanos a la línea recta. Este caso representa una curva de la experiencia de 72%. Esto es, el precio unitario promedio dis-

FIGURA 1-8
Precios de los circuitos integrados durante el periodo 1964-1972.

Fuente: Robert N. Noyce, "Microelectronics", en *Scientific American*, septiembre de 1977. © 1977 por Scientific American Inc. Derechos reservados. Reimpreso con autorización del editor.

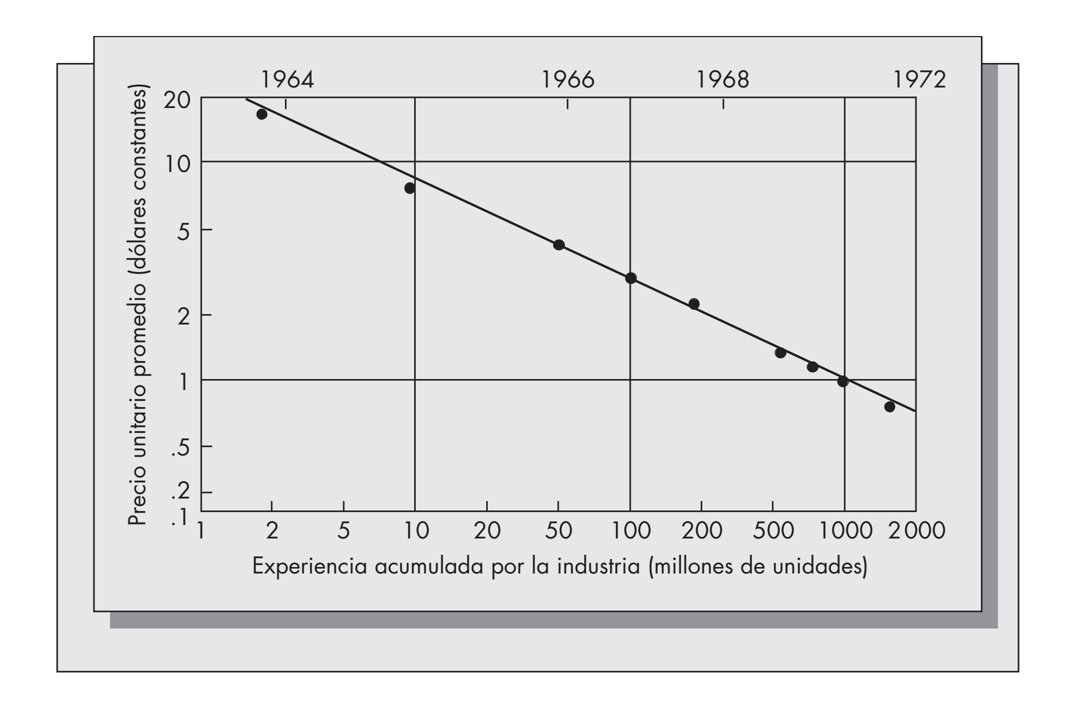
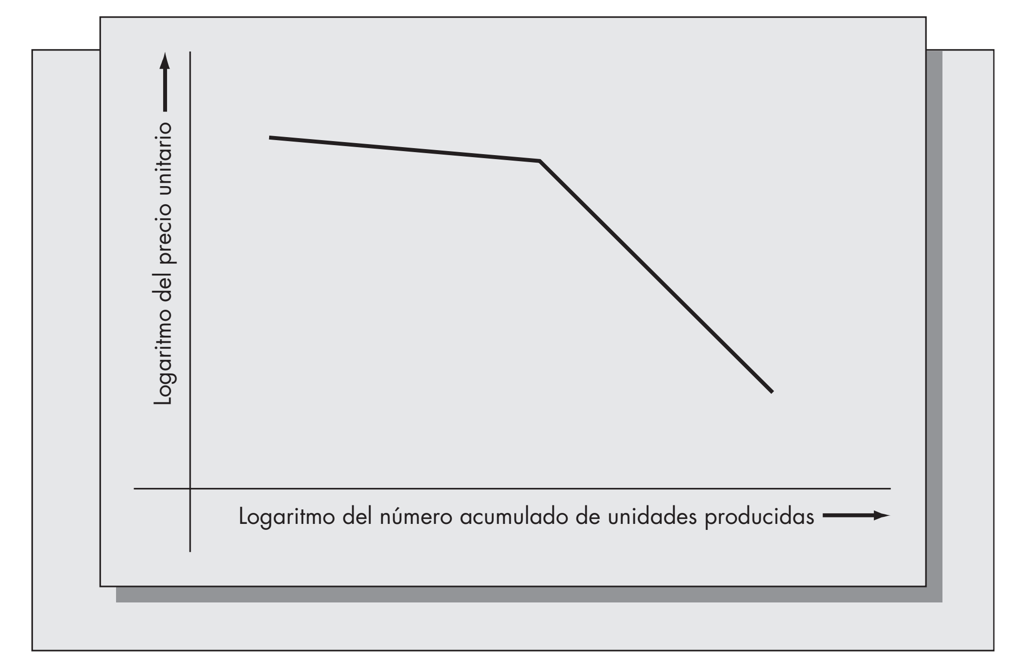


FIGURA 1-9
Curva "retorcida" de la experiencia debido a la fijación de precios rectores.



minuye en cerca de 72% de su valor anterior por cada duplicación de la producción acumulada de circuitos integrados por toda la industria.

Las curvas de la experiencia se miden generalmente en términos del costo unitario de producción. En la mayoría de las circunstancias, el comportamiento del precio de un producto o familia de productos sigue muy de cerca el del costo de producción. Sin embargo, en algunos casos se presentan precios rectores. Esto es, los precios permanecen prácticamente estables a lo largo de un periodo en el que los costos de producción disminuyen. Más tarde, cuando se intensifican las presiones competitivas del mercado, es necesario disminuir los precios con mayor rapidez que los costos hasta que aquéllos se equilibren. Ello puede provocar un retorcimiento en la curva de la experiencia si se está midiendo el precio en vez del costo contra el volumen acumulado. Este tipo de fenómeno aparece en la figura 1-9. Hayes y Wheelwright (1984, p. 243) dan dos ejemplos de precios rectores y su efecto sobre la curva de la experiencia. Las

curvas de la experiencia para el precio de los hornos de gas y el cloruro de polivinilo son ejemplos de este fenómeno.

Recientemente, las curvas de aprendizaje han estado sujetas a fuerte crítica en la literatura en diferentes aspectos: *a*) carecen de una justificación teórica; *b*) confunden los efectos del aprendizaje, economías de escala y otras mejoras tecnológicas, y *c*) se enfocan más en el costo que en las utilidades (Devinney, 1987). Sin embargo, es claro que dichas curvas son descriptores precisos de la forma en que los costos y las horas de mano de obra marginales disminuyen como una función de la experiencia acumulada obtenida por la compañía o la industria.

Curvas de experiencia y aprendizaje, y estrategia de manufactura

Definimos una **estrategia de la curva de aprendizaje** como una en que la meta principal es reducir los costos de producción siguiendo los lineamientos indicados por la curva de aprendizaje. Ford Motors adoptó una estrategia de la curva de aprendizaje al buscar reducir los costos del Modelo T durante el periodo de 1909 a 1923. Abernathy y Wayne (1974) mostraron que el precio de venta del Modelo T durante este periodo siguió muy de cerca una curva de la experiencia de 85%. La estrategia de Ford durante este tiempo estaba claramente dirigida a la reducción de costos; la compañía adquirió o construyó nuevas instalaciones, incluyendo altos hornos; molinos y aserraderos; un ferrocarril; talleres de costura; hornos de carbón; fábrica de papel; una planta de vidrio, y una planta de cemento. Esto permitió a Ford integrar verticalmente las operaciones, lo que dio como resultado menores tiempos de producción y menores inventarios, esto es, una estrategia similar a la de la filosofía justo a tiempo comentada anteriormente en este mismo capítulo.

Una estrategia de la curva de aprendizaje no necesariamente es la mejor opción para grandes horizontes de planeación. Abernathy y Wayne (1974) argumentaron que cuando la estrategia de manufactura se basa en la reducción de costos, se rigidiza la innovación. Ante el cambio en el gusto de los consumidores en la década de 1920, la atención de Ford a la reducción de costos y la estandarización del proceso de manufactura del Modelo T dio como resultado una notoria lentitud para adaptarse a los patrones cambiantes de las preferencias del cliente. La pérdida de Ford se convirtió en la ganancia de General Motors. GM respondió con rapidez a las necesidades del cliente, dándose cuenta de que el diseño abierto del Modelo T se volvería obsoleto muy pronto. Ford se vio entonces luchando para sobrevivir a finales de la década de 1930 después de haber disfrutado de un dominio prácticamente completo del mercado. Su supervivencia implicó romper con la anterior estrategia de curva del aprendizaje rígida, y cambiar a una basada en la innovación.

Otro ejemplo de una compañía que sufrió de una estrategia de curva de aprendizaje es Douglas Aircraft. El concepto de curva de aprendizaje estaba fuertemente enraizado en la industria aeronáutica. Douglas realizó varios compromisos en los años 1960 para la entrega de aeronaves basándose en la extrapolación de costos a lo largo de la curva de aprendizaje. Sin embargo, debido a cambios imprevistos en el diseño del producto, los costos fueron más elevados de lo previsto, y no pudo cumplir con los compromisos de entrega. Douglas se vio forzado a una fusión con McDonnell Company como resultado de los problemas financieros que experimentó.

No estamos sugiriendo con estos ejemplos que sea equivocada una estrategia de la curva de aprendizaje. La estandarización y la reducción de costos basadas en la producción de gran volumen han sido las claves del éxito para muchas compañías. Cuando no se logra un tiempo rápido para alcanzar un gran volumen se puede presentar un desastre en un mercado altamente competitivo. Lo que estamos diciendo es que la estrategia de la curva de aprendizaje debe balancearse con la flexibilidad suficiente como para permitir responder a los cambios en el mercado. La estandarización no debe rigidizar la innovación y la flexibilidad.

Problemas para la sección 1.10

25. ¿Cuáles son los factores que contribuyen al fenómeno curva de aprendizaje/curva de la experiencia?
26. ¿Qué es una "estrategia de la curva de aprendizaje"? Describa la forma en que esta estrategia llevó a Ford a alcanzar el éxito hasta mediados de los años 1920, y lo metió en problemas después de esa época.

27. ¿Cuáles son algunos de los errores que pueden presentarse cuando se utilizan las curvas de aprendizaje y las curvas de la experiencia para la predicción de costos? Refiérase a la experiencia de Douglas Aircraft.
28. Considere el ejemplo de XYZ Corporation presentado en esta sección. Si la curva de aprendizaje permanece precisa, ¿cuánto tiempo llevará producir la unidad número 100 000?
29. Una nueva firma ha registrado cuidadosamente el tiempo requerido para fabricar su producto: una válvula de desconexión en las tuberías de gasolina.

Número acumulado de unidades producidas	Número de horas requeridas para la siguiente unidad
50	3.3
100	2.2
400	1.0
600	0.8
1 000	0.5
10 000	0.2

a) Calcule los logaritmos de los números en cada columna (use logaritmos naturales).

b) Grafique ln(horas) contra ln(unidades acumuladas) y trace a ojo una línea recta que se ajuste a los datos. Utilizando su ajuste aproximado, estime *a* y *b*.

c) Utilice los resultados de la parte *b*), estime el tiempo requerido para producir la primera unidad y el porcentaje adecuado de la curva de aprendizaje que se ajuste para estos datos.

d) Repita las partes *b*) y *c*), pero use un ajuste de mínimos cuadrados exactos para los logaritmos calculados en la parte *a*).

30. Considere la curva de aprendizaje obtenida en el problema 29. ¿Cuánto tiempo se requerirá para producir la unidad número 100 000, asumiendo que la curva de aprendizaje siga siendo válida?
31. Considere la curva de experiencia graficada en la figura 1-10. ¿Cuál es el porcentaje de la curva de experiencia que parece estar representada?

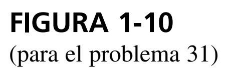
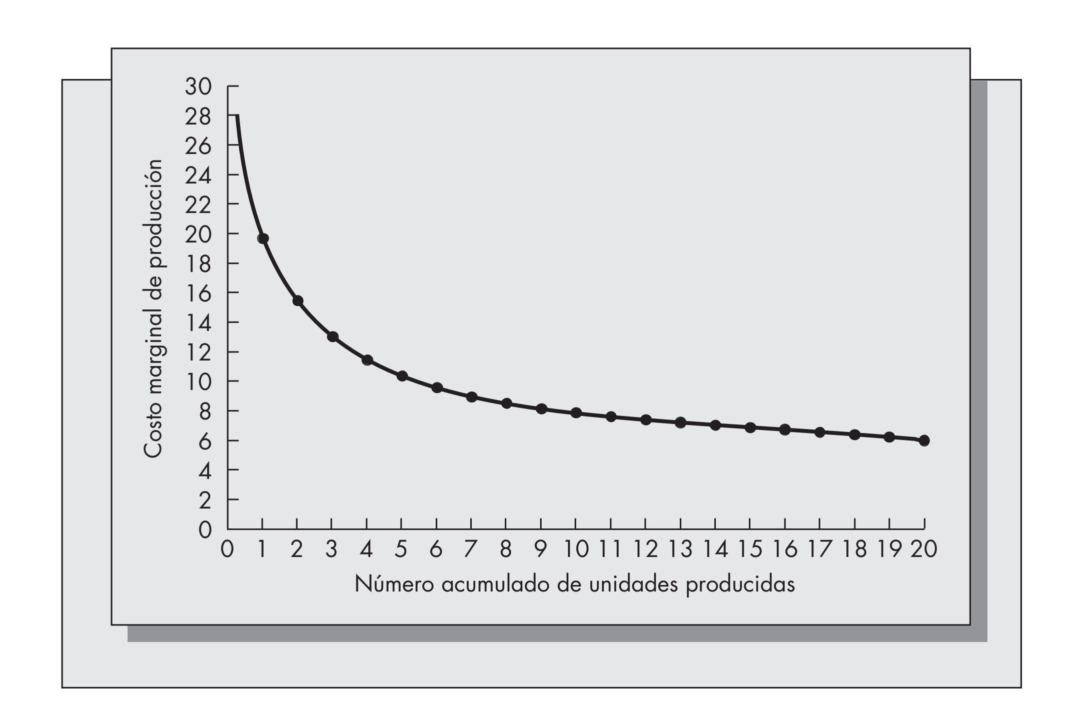


FIGURA 1-10
(para el problema 31)



32. Comente las limitaciones de las curvas de experiencia y aprendizaje.
33. Un analista predice que una curva de experiencia de 80% debe ser un predictor preciso para el costo de producir un nuevo producto. Supongamos que el costo de la primera unidad es de $1 000. ¿Qué costo predeciría el analista que tendría la producción de

    a) la unidad número 100?

    b) la unidad número 10 000?

1.11 PLANEACIÓN DEL CRECIMIENTO DE LA CAPACIDAD: UN PROBLEMA ESTRATÉGICO A LARGO PLAZO

La capacidad de una planta es el número de unidades que la planta puede producir durante un cierto tiempo. La política de capacidad juega un papel clave para determinar la posición competitiva de la compañía en el mercado. Una estrategia de capacidad debe tomar en cuenta una variedad de factores, incluyendo

- Patrones previstos de demanda.
- Costos de construir y operar nuevas instalaciones.
- Nueva tecnología de proceso.
- Estrategias de los competidores.

La capacidad de planeación es un asunto extremadamente complejo. Cada vez que una compañía considera expandir su capacidad productiva existente, debe buscar de entre un sinfín de posibilidades. En primer lugar, se debe decidir si se incrementa la capacidad mediante la modificación de las instalaciones existentes. Desde un punto de vista general, esta es una alternativa atractiva. Es más barato llevar a cabo cambios drásticos en procesos y plantas existentes que construir nuevas instalaciones. Sin embargo, una estrategia de este tipo finalmente puede resultar económica en lo que no tiene importancia y costosa en lo relevante. Existe evidencia sustancial de que las plantas focalizadas son mucho más productivas. Se suelen presentar muy rápidamente rendimientos decrecientes cuando una compañía trata de ir más allá del valor óptimo de la capacidad productiva de una cierta ubicación.

Dada la decisión de seguir adelante con la construcción de una nueva planta, es necesario resolver muchas interrogantes. Éstas incluyen:

1. *Cuándo.* El momento en el tiempo en que se realiza la construcción de nuevas instalaciones es una consideración importante. Los tiempos de demora para la construcción y los patrones cambiantes de demanda son dos factores que afectan la cuestión del tiempo.

2. *Dónde.* Ubicar nuevas instalaciones es un asunto complejo. La consideración de la logística de flujos de material sugiere que las nuevas instalaciones se ubiquen cerca de los proveedores de materias primas y de puntos de venta de mercado. Si los costos de mano de obra constituyen un aspecto clave, se puede optar por lugares en el extranjero. Los estados y los municipios otorgan algunas veces incentivos fiscales con el fin de atraer nueva industria. El costo de vida y la conveniencia geográfica son factores que afectarán la capacidad de la compañía para contratar y conservar empleados calificados.

3. *Cuánto.* Una vez que la dirección ha decidido cuándo y dónde agregar una nueva capacidad, debe decidirse el tamaño de las nuevas instalaciones. Agregar demasiada capacidad significa que ésta será subutilizada. Este es un problema especialmente serio cuando escasea el capital. Por otro lado, agregar menos capacidad de la necesaria significa que la compañía enfrentará en un futuro cercano el problema de tener que incrementar de nuevo la capacidad.

Economías de escala y economías de alcance

Las **economías de escala** por lo general se consideran las principales ventajas de expandir la capacidad existente. Panzer y Willig (1981) introdujeron el concepto de **economías de alcan-**

ce que definieron como los ahorros en los costos que se pueden obtener al combinar la producción de dos o más líneas de producto en un solo lugar. La idea es que los procesos de manufactura para estas líneas de producto pueden compartir parte del mismo equipo y personal, de manera que el costo de producción en un lugar pueda ser menor que el de dos o más lugares diferentes.

La noción de economías de alcance se extiende más allá de los ahorros en costos directos que la compañía puede obtener mediante la combinación de la producción de dos o más productos en un solo lugar. Muchas veces es necesario duplicar diversas funciones para atender diferentes lugares. Estas funciones incluyen sistemas de almacenamiento y recuperación de información y personal de soporte. Dichas actividades son más fáciles de coordinar si se encuentran en un mismo lugar. La compañía puede también lograr economías de alcance mediante la localización de diferentes instalaciones en una misma región geográfica. De esta forma los empleados pueden, en caso de ser necesario, recurrir al talento de personal clave en un lugar cercano.

Goldhar y Jelinek (1983) sostienen que las consideraciones de las economías de alcance apoyan la inversión en nueva tecnología de manufactura. Los sistemas flexibles de manufactura y la manufactura integrada por computadora dan como resultado "eficiencias derivadas de la variedad, no del volumen". Estos tipos de sistemas, dicen los autores, permiten que la compañía produzca de manera económica múltiples productos en lotes de tamaños pequeños, usando el mismo equipo multipropósito. (Los sistemas flexibles de manufactura se comentan con mayor detalle en el capítulo 10.)

La dirección debe sopesar los beneficios que la compañía puede obtener mediante la combinación de líneas de productos en un solo lugar contra las desventajas de la falta de focalización anteriormente comentadas. Demasiadas líneas de producto producidas en las mismas instalaciones pueden provocar que las diversas operaciones de manufactura interfieran unas con otras. El tamaño apropiado y la diversidad de las funciones de una sola planta deben balancearse de manera que la compañía pueda realizar economías de escala sin permitir que la planta pierda su focalización esencial.

Hacer o comprar: un problema prototipo de expansión de capacidad

Un problema clásico que enfrentan las compañías se conoce como la **decisión de hacer o comprar**. La compañía puede comprar el producto de una fuente externa a c_1 por unidad, pero puede producirlo internamente por un precio por unidad menor, $c_2 < c_1$. Sin embargo, para poder producir internamente el producto, la compañía debe invertir $\$K$ para expandir la capacidad de producción. ¿Qué estrategia debe adoptar la compañía?

El problema de hacer o comprar conlleva muchos de los elementos del problema de expansión de capacidad general. Deja en claro el dilema esencial entre la inversión y las economías de escala. El costo total para la compañía de producir x unidades es $K + c_2x$. Esto es equivalente a $K/x + c_2$ por unidad. Conforme x se incrementa, el costo por unidad de producción disminuye, ya que K/x es una función decreciente de x. El costo de comprar fuera es c_1 por unidad, independientemente de la cantidad ordenada. Al graficar los costos totales tanto de la producción interna como de la compra externa, podemos encontrar el punto en que se igualen los costos. Esto se conoce como la cantidad de equilibrio. Las curvas de equilibrio se muestran en la figura 1-11.

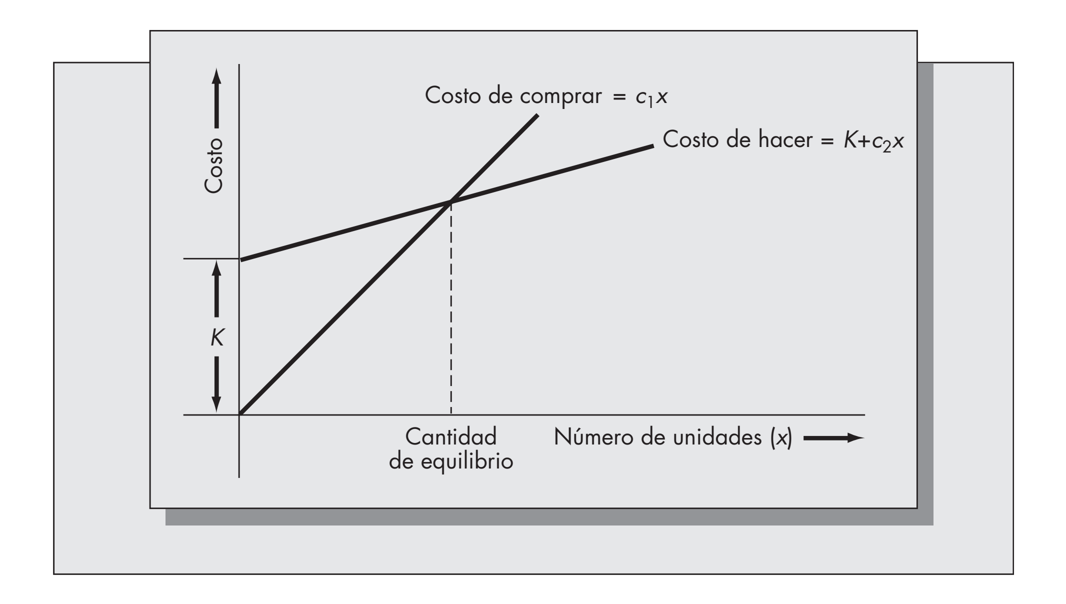
La cantidad de equilibrio resuelve

$$K + c_2x = c_1x,$$

resultando en $x = K/(c_1 - c_2)$.

Ejemplo *1.2* Un importante fabricante de computadoras a nivel internacional está diseñando un nuevo modelo de computadora personal y debe decidir si produce los teclados internamente o si los compra a un proveedor externo. El proveedor está dispuesto a vender los teclados en $50 cada uno, pero el fabricante estima que la compañía puede producirlos por $35 cada uno. La dirección estima que expandir la planta actual y comprar el equipo necesario para elaborar los teclados costaría $8 millones. ¿Deben llevar a cabo la expansión?

FIGURA 1-11
Curvas de equilibrio.



La cantidad de equilibrio es

$$x = 8\,000\,000/(50 - 35) = 533\,333.$$

Por tanto, la compañía tiene que vender cuando menos 533 333 teclados para poder justificar la inversión de $8 millones que se requiere para la expansión.

Este tipo de curvas de equilibrio resultan útiles para estimar rápidamente la conveniencia de una adición de capacidad. Su principal limitación es que son estáticas. No consideran los aspectos dinámicos del problema de capacidad, que no puede ignorarse en la mayoría de los casos. Éstos incluyen cambios en el patrón de demanda anticipado y consideraciones del valor del dinero en el tiempo. Aun como modelos estáticos, las curvas de equilibrio son meras aproximaciones. Ignoran los efectos del aprendizaje sobre la producción; esto es, que el costo marginal de producción disminuye con el incremento en el número de unidades producidas. (Las curvas de aprendizaje se comentan con detalle en la sección 1.10.) Dependiendo de la estructura de la función de producción, puede resultar económicamente viable producir algunas unidades internamente y comprar algunas unidades en el exterior. Manne (1967) comenta las implicaciones de algunas de estas cuestiones.

Política dinámica de expansión de la capacidad

Las decisiones de capacidad deben realizarse en un ambiente dinámico. En particular, la dinámica de un patrón cambiante de demanda determina cuándo debe invertir la compañía en nueva capacidad. Dos objetivos encontrados para la planeación de la capacidad son

1. Maximizar la participación de mercado
2. Maximizar la utilización de la capacidad

Una compañía que basa su estrategia a largo plazo en la maximización de la utilización de la capacidad corre el riesgo de incurrir en faltantes durante periodos con demanda mayor a la prevista. Una estrategia alternativa para incrementar la capacidad productiva consiste en producir para inventario y permitir que éste absorba las fluctuaciones de la demanda. Sin embargo, esto puede ser muy riesgoso. Los inventarios pueden volverse obsoletos, y los costos de mantenerlos pueden convertirse en una carga financiera.

Por otro lado, una compañía puede asumir la estrategia de mantener un "colchón de capacidad". Este colchón de capacidad consiste en un excedente de la capacidad que la compañía puede utilizar para responder a incrementos repentinos en la demanda; pone a la compañía en una posición adecuada para capturar una porción más grande del mercado si surge la oportunidad.

Considere el caso en que la demanda muestra una tendencia lineal creciente. Dos políticas (*a* y *b*) están representadas en la figura 1-12. En ambos casos, la compañía adquiere nueva capacidad en intervalos igualmente espaciados x, $2x$, $3x$, . . . , e incrementa la capacidad en una misma cantidad en esos tiempos. Sin embargo, en el caso *a*) la capacidad va por delante de la demanda, lo que significa que la firma mantiene un excedente de la capacidad en exceso en todo momento; mientras que en el caso *b*) la capacidad va por detrás de la demanda, lo que significa que la capacidad existente se utiliza plenamente en todo momento. Con el transcurso del tiempo, seguir la política *a*) o *b*) da como resultado el excedente en la capacidad (o la escasez de capacidad, si fuera el caso) que se muestra en la figura 1-13.

FIGURA 1-12
Estrategias de planeación de capacidad.

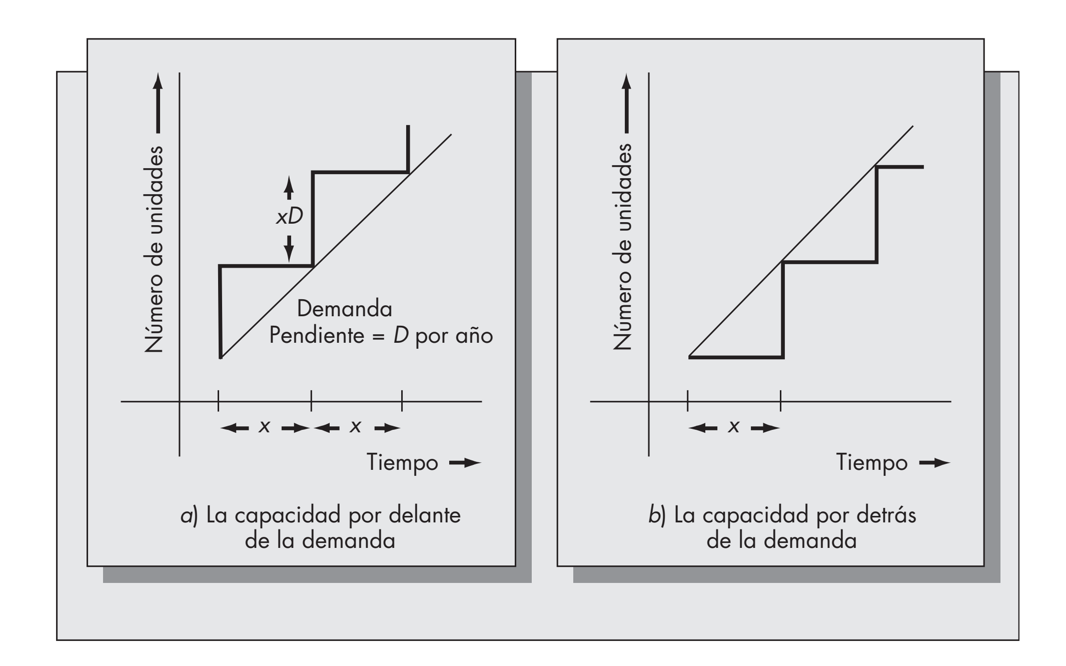
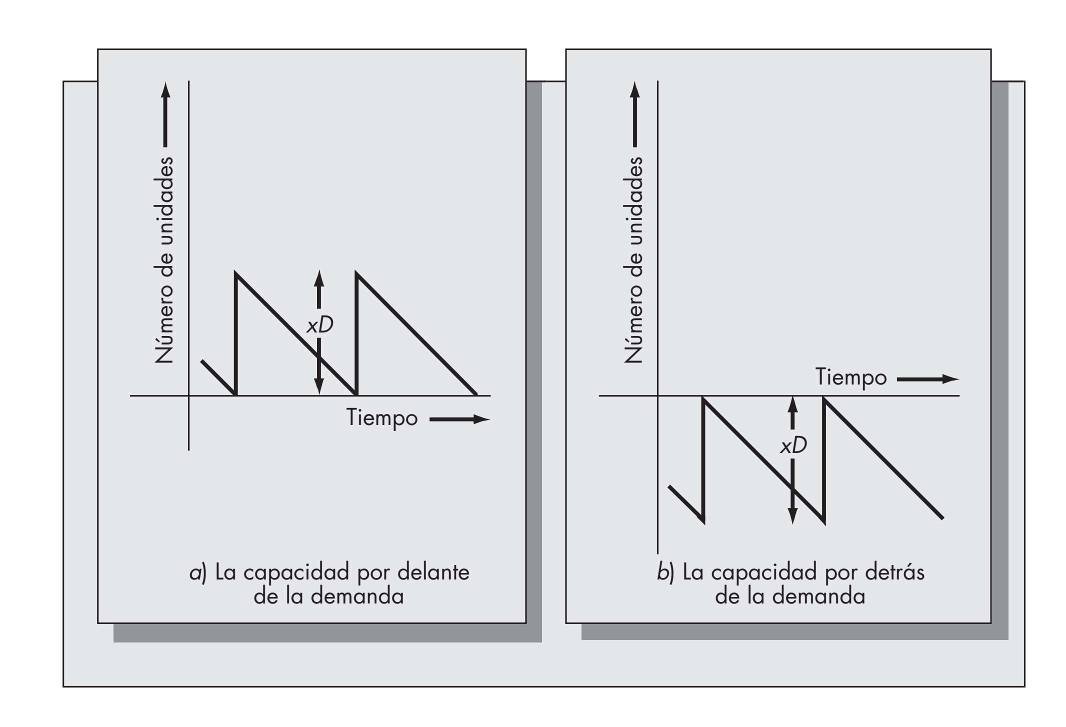


FIGURA 1-13
Trayectoria en el tiempo del excedente o deficiencia de la capacidad.



Considere el siguiente modelo particular que aparece en Manne (1967). Defina

D = Incremento anual en la demanda.
x = Intervalo de tiempo entre la introducción de plantas sucesivas.
r = Tasa de descuento anual, compuesta continuamente.
$f(y)$ = Costo de abrir una planta de capacidad y.

A partir de la figura 1-12*a* (que es la estrategia asumida en el modelo), vemos que si el intervalo de tiempo para el reemplazo de la planta es x, entonces el tamaño de la planta en cada reemplazo es xD. Más aún, el valor presente de un costo de \$1 incurrido t años en el futuro es dado por e^{-rt}. (En el apéndice 1-A se hace una presentación sobre el descuento y el valor del dinero en el tiempo.)

Defina $C(x)$ como la suma de los costos descontados para un horizonte infinito dando la apertura de una planta en el tiempo cero. Se desprende que

$$\begin{aligned} C(x) &= f(xD) + e^{-rx}f(xD) + e^{-2rx}f(xD) + \cdots \\ &= f(xD)[1 + e^{-rx} + (e^{-rx})^2 + (e^{-rx})^3 + \cdots] \\ &= \frac{f(xD)}{1 - e^{-rx}}. \end{aligned}$$

La experiencia ha mostrado que una representación de $f(y)$ que explica las economías de escala para plantas en diferentes industrias es

$$f(y) = ky^a,$$

donde k es una constante de proporcionalidad. El exponente a mide la razón entre los costos incrementales y los costos promedio de una unidad de capacidad de planta. Un valor de 0.6 parece ser común (conocido como la regla de las seis décimas). Mientras se cumpla que $a < 1$, hay economías de escala en la construcción de la planta, ya que para duplicar el tamaño de la planta se requiere un incremento inferior al doble en los costos de construcción. Para ver esto, considere la proporción

$$\frac{f(2y)}{f(y)} = \frac{k(2y)^a}{k(y)^a} = 2^a.$$

Al sustituir $a = 0.6$, se obtiene $2^a = 1.516$. Esto significa que si $a = 0.6$ es preciso, se puede duplicar la capacidad de la planta incrementando la inversión en dólares en alrededor de un 52%. De ahora en adelante, asumimos que $0 < a < 1$ para que haya economías de escala en el tamaño de la planta.

Para una determinada forma de $f(y)$, podemos determinar el momento óptimo de adiciones de planta y, por tanto, el tamaño óptimo de nuevas plantas. Si $f(y) = ky^a$, entonces

$$C(x) = \frac{k(xD)^a}{1 - e^{-rx}}.$$

Considere el logaritmo de $C(x)$:

$$\begin{aligned} \log[C(x)] &= \log[k(xD)^a] - \log[1 - e^{-rx}] \\ &= \log(k) + a\log(xD) - \log[1 - e^{-rx}]. \end{aligned}$$

Se puede demostrar que la función de costo $C(x)$ tiene un mínimo único con respecto a x y aún más, que el valor de x para el cual la derivada de $\log[C(x)]$ es cero es el valor de x que

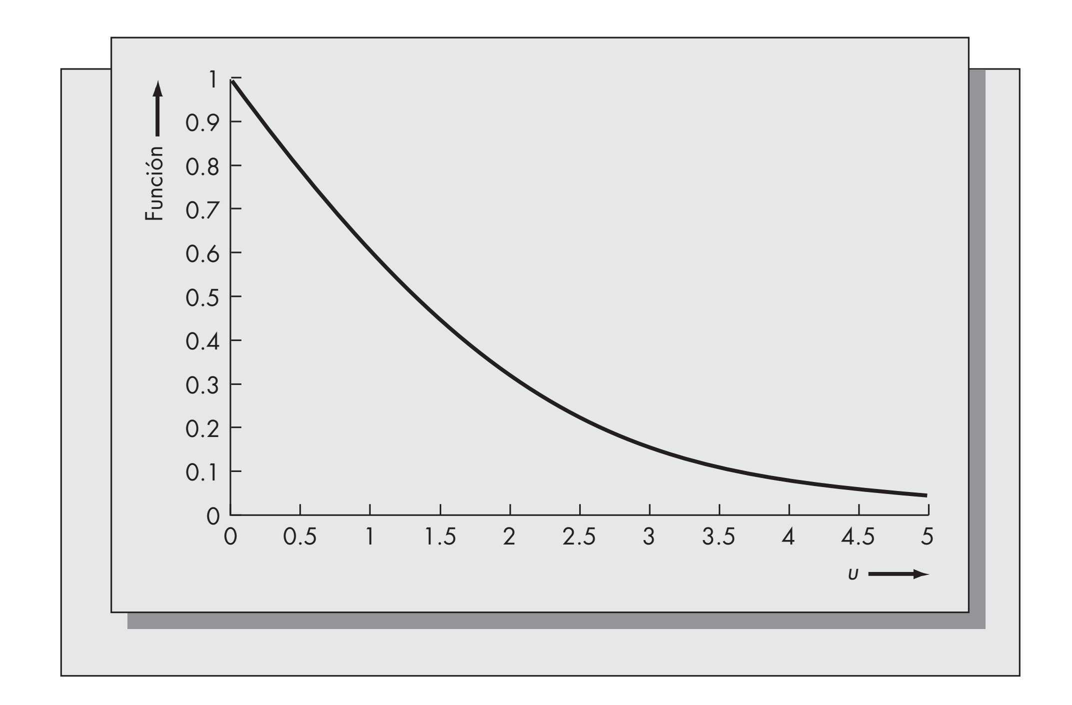


FIGURA 1.14
La función $u/(e^u - 1)$.

minimiza a $C(x)$. Es fácil mostrar[1] que la solución óptima satisface

$$\frac{rx}{e^{rx} - 1} = a.$$

La función $f(u) = u/(e^u - 1)$ se muestra en la figura 1-14, donde $u = rx$. Al localizar el valor de a sobre el eje de las ordenadas, se puede encontrar el valor óptimo de u en el eje de las abscisas.

Ejemplo *1.3*

Una compañía de productos químicos está planeando un incremento de capacidad de producción. La compañía ha estimado que el costo de agregar nueva capacidad obedece a la ley

$$f(y) = .0107y^{.62},$$

donde el costo se mide en millones de dólares y la capacidad en toneladas por año. Por ejemplo, al sustituir $y = 20\,000$ toneladas da un costo de planta de $f(y) = \$4.97$ millones. Es más, supongamos que la demanda está creciendo a una tasa constante de 5 000 toneladas por año y los costos futuros se descuentan con una tasa de interés de 16%. De la Figura 1-14 vemos que, si $a = .62$, el valor de u es aproximadamente de 0.9. Al resolver para x, obtenemos el tiempo óptimo de aperturas de plantas nuevas:

$$x = u/r = .9/.16 = 5.625 \text{ años}.$$

El valor óptimo de la capacidad de la planta debe ser $xD = (5.625)(5\,000) = 28\,125$ toneladas. Al sustituir $y = 28\,125$ en la ecuación para $f(y)$ se obtiene que el costo de cada planta en la solución óptima es de \$6.135 millones.

Mucha de la investigación sobre el problema de expansión de la capacidad consiste en extensiones de modelos de este tipo. Este modelo particular puede ser de ayuda en algunas

---

[1] $$\frac{d\log[C(x)]}{dx} = \frac{aD}{xD} - \frac{(-e^{-rx})(-r)}{1 - e^{-rx}} = \frac{a}{x} - \frac{r}{e^{rx} - 1} = 0,$$

que da $\frac{rx}{e^{rx} - 1} = a$.

circunstancias, pero ignora varias características fundamentales que uno esperaría encontrar en el mundo real:

1. *Vida finita de la planta.* El modelo supone que, una vez construida, una planta tiene una vida infinita. Sin embargo, las compañías cierran plantas por diversas razones: el equipo se vuelve obsoleto, o poco confiable, y no puede reemplazarse fácilmente. Los costos o los requerimientos de mano de obra pueden imponer un traslado hacia lugares menos costosos en el mismo país o a lugares en el extranjero. Los principales cambios en la tecnología de proceso pueden no ser fácilmente adaptables a las instalaciones existentes.

2. *Patrones de demanda.* Hemos asumido que la demanda crece a una tasa constante por año. Se han propuesto modelos para tomar en cuenta patrones de crecimiento de la demanda más complejos. En realidad, la incertidumbre de la demanda es un factor clave. En muchas industrias, la competencia extranjera ha hecho significativas incursiones en los mercados ya establecidos, forzando así a replantear las estrategias previas.

3. *Desarrollos tecnológicos.* El modelo asume que la capacidad de todas las nuevas plantas construidas permanece constante y que el costo de construir una planta de un tamaño dado permanece también constante. Esto es obviamente irrazonable. Con regularidad se presentan cambios drásticos en la tecnología de procesos, lo que modifica tanto el tamaño máximo de las nuevas plantas como los costos asociados para un cierto tamaño de una planta.

4. *Regulaciones gubernamentales.* Las restricciones ambientales y de seguridad pueden limitar las opciones de escala y ubicación de la planta.

5. *Costos indirectos*. La mayor parte de los modelos de ubicación y expansión no toman en cuenta explícitamente los costos indirectos. Durante la crisis energética de finales de los setenta, los costos de los energéticos se dispararon, ocasionando estragos en los presupuestos de costos indirectos de las plantas.

6. *Incentivos fiscales.* Las implicaciones financieras del tamaño y ubicación de nuevas instalaciones deben considerarse en el contexto de la planeación fiscal. Los autoridades locales o estatales suelen ofrecer incentivos fiscales a las grandes compañías que están considerando lugares alternativos para la construcción de nuevas instalaciones.

Una cuestión interesante es si los modelos de este tipo captan la forma en que realmente las compañías han tomado las decisiones de expansión en el pasado. Existe evidencia preliminar de que no lo hacen. Lieberman (1987) intentó evaluar los factores que motivaron a las compañías a construir nuevas plantas de productos químicos durante el periodo de 1957 a 1982 y descubrió que el tamaño de las nuevas plantas se incrementó aproximadamente 8% por año, independientemente de las condiciones de mercado. En periodos de gran demanda, las compañías construyeron más plantas. Este estudio preliminar indica que el pensamiento racional que lleva a modelos como el desarrollado en esta sección no refleja de manera precisa la forma en que las compañías toman decisiones respecto al tamaño de las plantas. Sus resultados sugieren que las compañías construyen las plantas lo más grandes posible con la tecnología existente (dada la existencia de las suficientes economías de escala, sin embargo, esta política puede resultar teóricamente óptima).

Problemas con la ubicación de una planta

Esta sección se relaciona con las decisiones respecto a la expansión de la capacidad, específicamente con determinar la cantidad y el tiempo de las nuevas adiciones a la capacidad. Una cuestión relacionada con esto es la **ubicación de las nuevas instalaciones**. Decidir dónde ubicar una planta es un problema complejo. Es necesario que la dirección considere cuidadosamente muchos factores antes de hacer la selección final.

La siguiente información acerca de la planta en sí es relevante para tomar la decisión de ubicación.

1. *Tamaño de la planta.* Esto incluye la extensión de terreno que se requiere, el número de metros cuadrados de espacio necesarios para la estructura de la edificación y las restricciones que pueden surgir como resultado de requerimientos especiales.
2. *Líneas de producto que se van a producir.*

3. *Tecnología de proceso por utilizar.*
4. *Requerimientos de la fuerza laboral.* Esto incluye tanto el número de trabajadores necesarios como la especificación de las habilidades particulares que se requieren.
5. *Necesidades de transportación.* Dependiendo de la naturaleza del producto producido y de los requerimientos de materias primas, es probable que la planta tenga que ubicarse cerca de carreteras importantes o de vías ferroviarias.
6. *Requerimientos de servicios públicos.* Éstos incluyen necesidades de energía, agua, drenaje o combustibles fósiles como el gas natural. Las plantas con grandes requerimientos energéticos deben ubicarse en áreas en las que la energía es menos costosa o cerca de fuentes de energía hidroeléctrica.
7. *Aspectos ambientales.* Debido a las regulaciones gubernamentales, habrá pocos lugares disponibles si la planta produce gran cantidad de productos de desecho.
8. *Interacción con otras plantas.* Si la planta es un satélite de instalaciones existentes, es probable que la dirección desee ubicar la nueva planta cerca de las otras.
9. *Consideraciones internacionales.* El hecho de que las nuevas instalaciones se ubiquen dentro del mismo país o en el extranjero es una cuestión muy sensible. Aunque los costos de mano de obra pueden ser menores en algunos sitios, como el Lejano Oriente, los aranceles, las cuotas de importación, los costos de inventarios en tránsito y la capacidad de respuesta al mercado también deben considerarse.
10. *Tratamiento fiscal.* Una variable importante son las consideraciones fiscales que se den en la decisión de la ubicación. Algunos países, como Irlanda, ofrecen un trato fiscal favorable para alentar el establecimiento de nuevas industrias. Existen también significativas diferencias en las leyes fiscales de los estados diseñadas para atraer a los fabricantes nacionales.

Los modelos matemáticos son útiles para ayudar a la toma de muchas decisiones operacionales. Sin embargo, tiene por lo general un valor limitado para determinar la ubicación adecuada de una nueva planta. Debido a que son muchos los factores y restricciones que se deben considerar en el proceso de decisión, dichas decisiones se toman con frecuencia con base en las aportaciones de una o más divisiones de la compañía, y el proceso de decisión puede llevar varios años. (En el capítulo 9 se explorarán técnicas matemáticas para tomar la decisión de la ubicación, ya que trata de la disposición y la ubicación de instalaciones.) Schmenner (1982) ha examinado el proceso de decisión en varias de las compañías de la lista de *Fortune 500*. Sus resultados mostraron que en la mayor parte de estas compañías la decisión sobre la ubicación de nuevas instalaciones la tomó la dirección corporativa de la empresa o el mismo CEO (siglas en inglés de *Chief Executive Officer*, o el ejecutivo del máximo nivel), aun cuando la necesidad de nuevas instalaciones se haya originado al nivel de división. El grado de autonomía en la toma de decisiones a nivel de división dependió del estilo de administración de la compañía.

Basado en una encuesta, Schmenner reportó que los principales factores que influyeron en la toma de decisiones de localización de nuevas instalaciones fueron los siguientes:

1. *Costos de mano de obra.* Ésta fue una de las principales preocupaciones de industrias como la de ropa, piel, mobiliario y electrodomésticos. Fue menos preocupante para las industrias de capital intensivo.
2. *Sindicalismo.* Un factor motivante para que una compañía considere la expansión de instalaciones existentes y no la construcción de nuevas instalaciones, es el potencial que se tenga para eliminar la influencia sindical en una nueva instalación. Una fuerza laboral fresca puede ser más difícil de organizar.
3. *Proximidad a los mercados.* Cuando los costos de transportación representan una parte importante del costo de los artículos vendidos, es esencial ubicar las nuevas plantas cerca de mercados existentes.
4. *Proximidad a proveedores y recursos.* La decisión acerca de dónde ubicar plantas en ciertas industrias se basa en la ubicación de recursos. Por ejemplo, las compañías que producen madera o productos de papel deben ubicarse cerca de bosques, mientras que las que producen alimentos procesados deben hacerlo cerca de instalaciones agrícolas.

5. *Proximidad a otras instalaciones.* Muchas compañías tienden a colocar las divisiones de manufactura y las instalaciones corporativas en la misma área geográfica. Por ejemplo, IBM se originó en el condado de Westchester, Nueva York, y ubicó muchas de sus divisiones en ese mismo estado. Al ubicar personal clave cerca unos de otros, la compañía ha sido capaz de realizar economías de alcance.
6. *Calidad de vida en la región*. Cuando no hay otras cuestiones que dicten la selección de un lugar, elegir un sitio que resulte atractivo para los empleados puede ayudar al reclutamiento de personal clave. Esto es especialmente cierto en el caso de industrias de alta tecnología que deben competir por trabajadores con habilidades particulares.

Problemas para la sección 1.11

34. Una nueva compañía, Macrotech, planea producir un dispositivo para traducir de clave Morse a un mensaje escrito en una computadora personal y enviar mensajes escritos en clave Morse a través de ondas de radio. El dispositivo es interesante principalmente para radioaficionados entusiastas. El presidente, Ron Lodel, estima que la inversión inicial podría ser de $30 000. El costo de producción de cada unidad es de $20 y se vende en $85.

    a) ¿Cuántas unidades debe vender para que la compañía recupere su inversión inicial?

    b) ¿Cuál es el ingreso total en el volumen de equilibrio?

    c) Si el precio se incrementará a $100 cada uno, encuentre el volumen de equilibrio.

35. Para el problema anterior, supongamos que se espera que las ventas sean de 100 unidades durante el primer año y se incremente a una tasa de 40% por año. ¿Cuántos años tomará recuperar la inversión inicial de $30 000? Asumamos que cada unidad se vende en $85.
36. Un productor nacional de carreolas, Pramble, compra las llantas a una compañía que se encuentra en el norte de Inglaterra. Actualmente las llantas cuestan $4 cada una, pero por diferentes razones este precio se va a duplicar. Con el fin de producir las llantas ellos mismos, Pramble tendría que agregar capacidad a las instalaciones existentes a un costo de $800 000. Se calcula que el costo unitario de producción podría ser de $3.50. Actualmente, la compañía vende 10 000 carreolas al año. (Asumamos que son cuatro llantas por carreola.)

    a) Con la tasa actual de ventas, ¿cuánto tiempo llevaría recuperar la inversión requerida para la expansión?

    b) Si se espera que las ventas se incrementen a una tasa de 15% al año, ¿cuánto tiempo llevará recuperar la expansión?

37. Con base en experiencias pasadas, una compañía de productos químicos estima que el costo de nuevas adiciones de capacidad obedece a la ley

    $$f(y) = .0205y^{.58}$$

    donde y se mide en toneladas por año y $f(y)$ en millones de dólares. La demanda está creciendo a una tasa de 3 000 toneladas por año y el departamento de contabilidad recomienda una tasa de 12% por año para el descuento de costos futuros.

    a) Determine el tiempo óptimo para las adiciones de la planta y el tamaño óptimo de cada adición.

    b) ¿Cuál es el costo de cada adición?

    c) ¿Cuál es el valor presente del costo de las siguientes cuatro adiciones? Para fines de cálculo asuma que se acaba de realizar una adición. (Consulte el apéndice 1-A para ver una discusión sobre el descuento de costos.)

38. Una importante compañía petrolera está considerando el momento óptimo para la construcción de nuevas refinerías. Por experiencias pasadas, cada duplicación en el tamaño de una refinería en una ubicación da como resultado un incremento en los costos de construcción de cerca de 68%. Adicionalmente, el tamaño de una planta de 10 000 barriles por día cuesta $6 millones. Asuma que la demanda de petróleo está aumentando a una tasa constante de dos millones de barriles por año y que la tasa de descuento para costos futuros es de 15%.

    a) Encuentre los valores de k y a asumiendo una relación de la forma $f(y) = ky^a$. Asuma que y está en unidades de barriles por día.

    b) Determine el tiempo óptimo de las adiciones a la planta y el tamaño óptimo de cada planta.

    c) Suponga que la refinería más grande que se puede construir con la tecnología actual es de 15 000 barriles por día. Determine el tiempo óptimo de las adiciones a la planta y el tamaño óptimo de cada planta en este caso. (Asuma 365 días por año para sus cálculos.)

39. Comente los pros y contras de ubicar instalaciones de manufactura en el extranjero.

1.12 Resumen

Este capítulo trató sobre la importancia de la **estrategia de operaciones** y su relación con la estrategia general de negocios de la compañía. Las operaciones continúan creciendo en importancia a nivel compañía. Si bien una porción más bien grande de la manufactura directa continúa yéndose al extranjero, no debe subestimarse la importancia de la función de manufactura. El éxito de la estrategia de operaciones se puede medir en varias dimensiones. Éstas incluyen las medidas obvias de costo y características de producto, pero también la calidad, velocidad de entrega, confiabilidad de la entrega y flexibilidad.

El punto de vista clásico de la estrategia de manufactura, debido principalmente a Wickham Skinner, considera las siguientes cuatro dimensiones de estrategia: **horizonte de tiempo**, **focalización**, **evaluación** y **consistencia**. Diferentes tipos de decisiones se relacionan con diferentes horizontes de tiempo. Una planta debe diseñarse con un enfoque específico en mente, sea para minimizar el costo unitario o para maximizar la calidad del producto. Pueden aplicarse diversos criterios de evaluación para analizar la efectividad de una estrategia.

Escuchamos cada vez con más frecuencia que somos parte de una comunidad global. Cuando compramos productos hoy en día, nos preocupa menos el país de origen que las características del producto. ¿Cuántos consumidores de teléfonos celulares están siquiera conscientes de que las oficinas generales de Nokia se encuentran en Finlandia, Ericsson en Suecia y Motorola en Estados Unidos? Una cuestión interesante explorada por Michael Porter es: ¿por qué algunas industrias parecen florecer en algunos países? Aunque la respuesta es compleja, Porter sugiere que los cuatro factores que presentamos a continuación son de gran importancia: *condiciones de factor; condiciones de demanda; industrias de soporte y relacionadas, y estrategia, estructura y rivalidad de la compañía.*

Puede resultar difícil cambiar la forma en que una persona hace las cosas. Aún más difícil es cambiar la forma en que toda una compañía hace las cosas. Por esa razón, la **reingeniería de procesos de negocios (BPR)** es un proceso doloroso, aun cuando funcione. Los éxitos más importantes de la BPR han provenido de funciones de servicio, pero el concepto puede aplicarse a cualquier contexto. Es el proceso de repensar el cómo y el por qué se hacen las cosas de una cierta manera. Hecha de manera inteligente, la BPR puede lograr mejoras espectaculares. Sin embargo, también puede ser un proceso que consuma mucho tiempo y dinero.

El **justo a tiempo (JIT)** es una filosofía que surgió del sistema kanban desarrollado por Toyota. En el centro del enfoque está la eliminación de desperdicio. Se instalan sistemas para reducir los flujos de material a lotes pequeños para evitar grandes acumulaciones de inventarios de trabajo en proceso. Aunque JIT se ha desarrollado en la fábrica, es un concepto que ha sido aplicado también a la función de compras. La aplicación exitosa de las compras JIT

requiere del desarrollo de relaciones a largo plazo con los proveedores, y usualmente requiere de una gran proximidad con ellos. La mecánica de JIT se comenta con mayor detalle en el capítulo 7.

Ser capaces de entrar al mercado rápidamente con productos que la gente desea en los volúmenes que el mercado requiere es crucial si deseamos ser líderes en el mercado. La **competencia basada en el tiempo** significa que debe reducirse el tiempo que existe entre la concepción del producto y su aparición en el mercado. Para lograrlo, debemos realizar de manera concurrente el mayor número de tareas posible. En muchos casos, el tiempo para llegar al mercado es menos relevante que el tiempo para alcanzar un cierto volumen. Ser el primero en el mercado puede no significar mucho si no podemos satisfacer la demanda del producto.

Los espectaculares éxitos de los japoneses durante los setenta y ochenta se debieron en gran medida a la excelente **calidad** de sus productos manufacturados. Dos estadounidenses, Deming y Juran, visitaron Japón a principios de los cincuenta y jugaron un papel determinante en concientizar a los japoneses sobre la importancia de producir productos de calidad. El movimiento de calidad en Estados Unidos ha dado como resultado una mayor conciencia de la importancia de la calidad, reconocimiento a los logros mayores en esta área por medio del premio Malcom Baldridge e iniciación de programas relevantes como los círculos de calidad y el programa 6-sigma en Motorola. (En el capítulo 11 se discuten a detalle cuestiones estadísticas y organizaciones concernientes a la calidad.)

Es importante comprender los **ciclos de vida del proceso y del producto**. Ambos pasan por las cuatro etapas de inicio, crecimiento rápido, maduración y estabilización o declive. Es igualmente valioso también comprender qué tipos de procesos son apropiados para qué tipo de productos y servicios. Para este fin, Hayes y Wheelwright han desarrollado el concepto de la matriz producto-proceso.

Las **curvas de experiencia y de aprendizaje** son útiles para modelar la reducción en las horas de mano de obra o de los costos de un producto conforme se adquiere experiencia en la producción de un artículo o familia de artículos. Estas curvas han mostrado que obedecen a una ley exponencial y pueden ser predictores útiles del costo y tiempo requeridos para la producción. (La ley de Moore, debida a Gordon Moore, el fundador de Intel, predijo la duplicación del desempeño del chip cada 18 meses.) Esto es un ejemplo de una curva de experiencia, y la predicción ha continuado siendo exacta hasta nuestros días.

Comentamos dos métodos para ayudar a las decisiones de **expansión de capacidad**. Las curvas de equilibrio proporcionan un medio para determinar el volumen de ventas necesario para justificar la inversión en instalaciones nuevas o existentes. Se presenta un modelo sencillo para una política dinámica de expansión que proporciona el tiempo y el tamaño óptimos para nuevas instalaciones suponiendo un crecimiento constante de la demanda y el descuento de costos futuros. También comentamos cuestiones que surgen al tratar de decidir dónde ubicar nuevas instalaciones. Este problema es muy complejo debido a que existen muchos factores que se relacionan con la decisión de dónde ubicar instalaciones de administración, diseño y manufactura.

Problemas adicionales para el capítulo 1

40. ¿Qué es una estrategia de producción y de operaciones? Comente los elementos que tiene en común con las estrategias financieras y de mercadotecnia, y los elementos que las diferencian.
41. ¿Cuál es la diferencia entre el ciclo de vida del producto y el del proceso?, ¿en qué forma se relacionan estos conceptos?
42. Supongamos que la Corporación Mendenhall, productora de bolsas para dama, ha determinado que 73% de una curva de experiencia describe de manera precisa la evolución de sus costos de producción para una nueva línea. Si la primera unidad cuesta $100 producirse, ¿cuál será el costo de la unidad número 10 000 basándonos en la curva de experiencia?

43. La tienda departamental Delon's vende varias marcas de ropa propias además de varias marcas de diseñadores conocidos. Delon's está considerando construir una planta en Malasia para producir corbatas de seda. La planta tendría un costo para la compañía de \$5.5 millones. La planta podrá producir las corbatas en \$1.20 cada una. Por otro lado, Delon's puede mandar maquilar (subcontratar) la producción de las corbatas a un precio de \$3.00. ¿Cuántas corbatas tiene que vender Delon's en el mundo para equilibrar su inversión en la nueva planta?

44. Un fabricante de acero japonés está considerando expandir sus operaciones. Con base en su experiencia, calcula que las adiciones de capacidad obedecen la ley

$$f(y) = .00345y^{.51},$$

donde el costo $f(y)$ se mide en millones de dólares y y se mide en toneladas de acero producidas. Si la demanda de acero está creciendo a una tasa constante de 8 000 toneladas por año y los costos futuros se están descontando usando una tasa de interés de 10%, ¿cuál es el número óptimo de años entre las aperturas de nuevas plantas?

Los problemas siguientes están diseñados para ser resueltos en una hoja de cálculo.



45. Considere el siguiente problema de equilibrio: el costo de producir Q unidades, $c(Q)$, se describe mediante la curva

$$c(Q) = 48Q[1 - \exp(-.08Q)],$$

donde Q está en cientos de unidades de artículos producidos y $c(Q)$ está en miles de dólares.

a) Grafique la función $c(Q)$. ¿Qué forma tiene? ¿Qué fenómeno económico da origen a una curva de costo acumulada con esta forma?

b) ¿Para qué nivel de producción el costo acumulado de producción será igual a \$1 000 000?

c) Suponga que estas unidades pueden adquirirse de un proveedor externo a un costo de \$800 cada uno, pero la compañía debe invertir \$850 000 para construir instalaciones que sean capaces de producir estas unidades a un costo de $c(Q)$. ¿Cuál es el volumen acumulado de producción para el cual tiene sentido invertir en las instalaciones?



46. Se espera que los costos de mantenimiento de nuevas instalaciones sean de \$112 000 durante el primer año de operaciones. Se anticipa que estos costos se incrementarán a una tasa de 8% al año. Asumiendo una tasa de retorno de 10%, ¿cuál es el valor presente del flujo de costos de mantenimiento para los próximos 30 años?



47. Suponga que el proveedor de teclados descrito en el ejemplo 1.2 quiere ofrecer el siguiente esquema de descuento incremental por cantidad:

Costo por teclado	Cantidad del pedido
\$50	$Q \leq 100\,000$
\$45	$100\,000 < Q \leq 500\,000$
\$40	$500\,000 < Q$

Determine el costo para la compañía de ordenar cantidades en incrementos de 20 000 para $Q = 200\,000$ a $Q = 1\,000\,000$, y compárelo con el costo que tendría para la compañía producir internamente para los mismos valores de Q. ¿Cuál es la cantidad de equilibrio para los pedidos?

Apéndice 1-A

Cálculos del valor presente

Es muy común considerar el valor del dinero en el tiempo en el proceso de decisión al momento de considerar estrategias alternativas de inversión. La idea es que un dólar recibido hoy tiene mayor valor que uno recibido dentro de un año. Por ejemplo, si un dólar fuera depositado en una cuenta de ahorros que paga 5% de interés, tendrá un valor de $1.05 en un año. En términos más generales, si se invirtiera a una tasa de retorno de r (expresada en fracción), tendría un valor de $1 + r$ en un año, $(1 + r)^2$ en dos años, etcétera.

De la misma forma, un costo de $1 incurrido en cierto año tiene un valor presente inferior al de $1 hoy en día. Por ejemplo, a 5% de interés, ¿cuánto necesitaría una persona depositar en una cuenta para que la suma del principal más el interés sea igual a $1 dentro de un año? La respuesta es $1/(1.05) = 0.9524$. De manera similar, el valor presente de un costo de $1 incurrido en dos años a 5% es $1/(1.05)^2 = 0.9070$. En general, el valor presente de un costo de $1 incurrido dentro de t años asumiendo una tasa de retorno r es $(1 + r)^{-t}$.

Estos cálculos asumen que no hay composición. Composición significa que uno obtiene interés sobre el interés, por así decirlo. Por ejemplo, 5% compuesto semestralmente significa que uno gana 2.5% sobre $1 después de seis meses y 2.5% en el $1 original más el interés ganado en los primeros seis meses. Por tanto, el retorno total es

$$(1.025)(1.025) = \$1.050625$$

después de un año, o ligeramente más de 5%. Si el interés fuera compuesto trimestralmente, el dólar valdría

$$(1 + .05/4)^4 = 1.0509453$$

al final del año. La extensión lógica de esta idea es la composición continua. Un dólar invertido a 5% compuesto continuamente valdría

$$\begin{aligned} \lim_{n\to\infty} (1 + .05/n)^n &= e^{.05} \\ &= 1.05127 \end{aligned}$$

al final del año. El número $e = 2.7172818\ldots$ se define como

$$e = \lim_{n\to\infty} (1 + 1/n)^n.$$

Observe que la composición continua sólo incrementa la tasa de interés simple efectiva de 5 a 5.127%.

En términos más generales, C, invertido a una tasa de r durante t años compuesto continuamente vale Ce^{rt} al final de t años.

Invirtiendo el argumento, el valor presente de un costo de C incurrido dentro de t años asumiendo una composición continua a una tasa de descuento r es Ce^{-rt}. Un flujo de costos C_1, $C_2, \ldots, C_n$ incurrido en los tiempos $t_1, t_2, \ldots, t_n$ tiene un valor presente

$$\sum_{i=1}^{n} C_i e^{-rt_i}.$$

Un tratamiento amplio sobre descuento y su relación con el problema de la expansión de capacidad se puede encontrar en Freidenfelds (1981).

Bibliografía

Abernathy, W. J., y P. L. Townsend. "Technology, Productivity, and Process Change". *Technological Forecasting and Social Change 7*, 1975, pp. 379-396.

Abernathy, W. J. y K. Wayne. "Limits of the Learning Curve". *Harvard Business Review* 52, septiembre-octubre de 1974, pp. 109-119.

Becker, G. S. Citado en *Business Week*, 27 de enero de 1986, p. 12.

Bell, D. *The Coming of the Post Industrial Society: A Venture in Social Forecasting*. Nueva York: Basic Books, 1976.

Blackburn, J. D. *Time-Based Competition: The Next Battleground in American Manufacturing*. Nueva York: McGraw-Hill/Irwin, 1991.

Buffa, E. S. y R. K. Sarin. *Modern Production/Operations Management.* 8a. edición. Nueva York: John Wiley & Sons, 1987.

Business Week, 22 de octubre de 1990, pp. 94-95.

Business Week, "The Quality Imperative", 25 de octubre de 1991.

Cohen, S. S. y J. Zysman. *Manufacturing Matters: The Myth of the Post-Industrial Economy.* Nueva York: Basic Books, 1987.

Davis, D. "Beating the Clock". *Electronic Business*, mayo de 1989.

Dertouzes, M. L., R. K. Lester y R. M. Solow. *Made in America: Regaining the Productive Edge.* Cambridge, MA: MIT Press, 1989.

Devinney, T. M. "Entry and Learning". *Management Science* 33, 1987, pp. 706-724.

Dhalla, N. K. y S. Yuspeh. "Forget about the Product Life Cycle Concept". *Harvard Business Review* 54, enero-febrero de 1976, pp. 102-112.

Drucker, P. F. "Japan: New Strategies for a New Reality". *The Wall Street Journal*, 2 de octubre de 1991.

Fallon, M. *San Jose Mercury News*, 30 de septiembre de 1985, p. 11D.

Fine, C. H. y A. C. Hax. "Manufacturing Strategy: A Methodology and an Illustration". *Interfaces* 15, 1985, pp. 27-47.

Freeland, J. R. "A Survey of Just-in-Time Purchasing Practices in the United States". *Production and Inventory Management Journal*, segundo trimestre de 1991, pp. 43-50.

Freidenfelds, J. *Capacity Expansion: Analysis of Simple Models with Applications.* Nueva York: Elsevier North Holland, 1981.

Goldhar, J. P. y M. Jelinek. "Plan for Economies of Scope". *Harvard Business Review* 61, 1983, pp. 141-148.

Hammer, M. S. y J. Champy. *Reengineering the Corporation: A Manifesto for Business Revolution.* Nueva York: Harper Business, 1993.

Hammer, M. S. y S.A. Stanton. *The Reengineering Revolution.* Nueva York: Harper Business, 1995.

Hayes, R. H. y S. Wheelwright. "Link Manufacturing Process and Product Life Cycles". *Harvard Business Review* 57, enero-febrero de 1979, pp. 133-140.

Hayes, R. H. y S. Wheelwright. *Restoring Our Competitive Edge: Competing through Manufacturing*. Nueva York: John Wiley & Sons, 1984.

Hill, T. J. *Manufacturing Strategy—Text and Cases*. Nueva York: McGraw-Hill/Irwin, 1993.

Jones, P. y L. Kahaner. *Say It and Live It. The Fifty Corporate Mission Statements That Hit the Mark.* Nueva York: Currency Doubleday, 1995.

Krugman, P. *Peddling Prosperity: Economic Sense and Nonsense in the Age of Diminished Expectations.* Nueva York: W. W. Norton and Company, 1994.

Lieberman, M. "Scale Economies, Factory Focus, and Optimal Plant Size". Artículo presentado en la Escuela de Posgrados de Negocios de Stanford, 22 de julio de 1987.

Manne, A. S., editores. *Investments for Capacity Expansion: Size, Location, and Time Phasing*. Cambridge, MA: MIT Press, 1967.
Noyce, R. N. "Microelectronics". *Scientific American*, septiembre de 1977.
Panzer, J. C. y R. O. Willig. "Economies of Scope". *American Economic Review*, 1981, pp. 268-272.
Port, O. "Customers Move into the Driver's Seat". *Business Week*, 4 de octubre de 1999, pp. 103-106.
Porter, M. E. *Competitive Strategy: Techniques for Analyzing Industries and Competitors.* Nueva York: The Free Press, 1980.
Porter, M. E. *The Competitive Advantage of Nations.* Nueva York: The Free Press, 1990.
Prestowitz, C. V., Jr. *Trading Places.* Nueva York: Basic Books, 1989.
Schmenner, R. W. *Making Business Location Decisions*. Englewood Cliffs, NJ: Prentice-Hall, 1982.
Skinner, W. "The Focused Factory". *Harvard Business Review* 52, mayo-junio de 1974, pp. 113-121.
Skinner, W. *Manufacturing in the Corporate Strategy.* Nueva York: John Wiley & Sons, 1978.
Stasey, R. y C. J. McNair. *Crossroads: A JIT Success Story.* Nueva York: McGraw-Hill/Irwin, 1990.
Stewart, T. A. "The New American Century: Where We Stand". *Fortune*, primavera-verano 1991.
Wills, F. "Bully for Taurus". *Business Month*, febrero de 1990, p. 13.
Womack, J. P., D. T. Jones y D. Roos. *The Machine That Changed the World.* Nueva York: Harper Perennial, 1990.
World Economic Outlook, publicado por el Fondo Monetario Internacional. Washington, DC: marzo de 2003.

Capítulo **Dos**

Pronósticos

Panorama general del capítulo

Objetivo

Presentar e ilustrar los métodos más importantes para pronosticar la demanda en el contexto de la planeación de operaciones.

Puntos clave

1. *Características de los pronósticos*
   - Casi siempre estarán equivocados.
   - Un buen pronóstico también da una medida de error.
   - Pronosticar unidades en conjunto es más fácil que pronosticar unidades individuales.
   - Entre más a futuro se realizan los pronósticos, menos exactos son.
   - Una técnica de pronóstico no debe usarse para excluir información conocida.
2. *Pronósticos subjetivos.* Se refiere a los métodos que miden la opinión, ya sea individual o grupalmente. Entre los más conocidos se encuentran:
   - Agregados de la fuerza de ventas.
   - Encuestas al cliente.
   - Juicio de opinión ejecutiva.
   - El método Delphi.
3. *Métodos de pronósticos objetivos (métodos con series de tiempo y regresión).* Al utilizar los métodos de pronóstico objetivo, se realizan pronósticos con base en datos históricos. Los pronósticos con series de tiempo utilizan sólo la historia de la serie que se va a pronosticar, mientras que los modelos de regresión muchas veces incorporan el historial de otras series. En los pronósticos con series de tiempo, el objetivo es encontrar patrones predecibles y repetibles en los datos pasados. Con base en el patrón que se identifique, pueden usarse varios métodos. Los métodos con series de tiempo tienen la ventaja de que pueden incorporarse fácilmente en un programa de computadora para la actualización y el pronóstico automático. Entre los patrones repetibles estarían la tendencia lineal creciente o decreciente, la tendencia curvilínea (incluyendo el crecimiento exponencial), y las fluctuaciones estacionales. Cuando usamos la regresión, construimos un modelo causal que predice un fenómeno (la variable dependiente) con base en la evolución de uno o más fenómenos distintos (las variables independientes). Un ejemplo puede ser predecir el inicio o el final de una recesión tomando como base los proyectos de vivienda, que se consideran un indicador significativo de la salud de la economía.

4. *Evaluación de los métodos de pronóstico.* El error de pronóstico en cualquier periodo, e_t, es la diferencia entre el pronóstico para el periodo t, y el valor real de la serie realizado para el periodo t ($e_t = F_t - D_t$). Tres medidas comunes del error de pronóstico son la DAM (desviación absoluta media durante n periodos), el ECM (el error cuadrático medio durante n periodos), y el EPAM (el error porcentual absoluto medio durante n periodos).
5. *Métodos para pronosticar series de tiempo estacionario.* Consideramos dos métodos de pronóstico cuando el patrón subyacente de la serie resulta estacionario a lo largo del tiempo: los promedios móviles y el suavizamiento exponencial. Un *promedio móvil* es simplemente el promedio aritmético de las N observaciones más recientes. Los *pronósticos de suavizamiento exponencial* se basan en el promedio ponderado de la observación más reciente y en el pronóstico anterior. El peso aplicado a la observación más reciente es α, donde $0 < \alpha < 1$ y, el peso aplicado al último pronóstico es $1 - \alpha$. Ambos métodos se utilizan comúnmente en la práctica, pero el suavizamiento exponencial se prefiere en las aplicaciones de control de inventarios —especialmente en los sistemas grandes— porque no necesita almacenar tantos datos como los promedios móviles.
6. *Métodos para pronosticar series con tendencia.* Cuando existe una tendencia lineal ascendente o descendente en los datos, dos métodos comunes de pronóstico son la *regresión lineal* y el suavizamiento exponencial doble por medio del *método de Holt*. La regresión lineal se utiliza para ajustar una línea recta a datos pasados con base en el método de mínimos cuadrados, y el método de Holt utiliza ecuaciones separadas de suavizamiento exponencial para pronosticar la intercepción y la pendiente de las series de cada periodo.
7. *Métodos para pronosticar series estacionales.* Una serie de tiempo estacional es aquella que tiene un patrón de repetición regular durante el mismo periodo de tiempo. Por lo general, el margen de tiempo puede ser de un año, y los periodos pueden ser semanas o meses. El enfoque más sencillo para pronosticar series estacionales se basa en los factores estacionales multiplicativos. Un factor estacional multiplicativo es un número que indica el valor relativo de una serie en cualquier periodo comparado con el valor promedio a lo largo de un año. Supongamos que una estación consiste en 12 meses. Un factor estacional de 1.25 para un mes dado significa que la demanda en ese mes es 25% más alta que la demanda media mensual. El *método de Winters* es un método más complejo basado en un suavizamiento exponencial triple. Se utilizan tres ecuaciones de suavizamiento distintas para pronosticar la intercepción, la pendiente y los factores estacionales en cada periodo.
8. *Otras consideraciones.* El pronóstico económico y el análisis de series de tiempo son áreas de investigación muy ricas. Cuando se tiene un largo historial de datos, puede disponerse de métodos mucho más sofisticados. Los llamados Box-Jenkins se basan en la evaluación y el examen de la función de autocorrelación para determinar un modelo adecuado. La teoría del filtrado, originalmente desarrollada en el contexto de problemas en comunicaciones, puede muchas veces adaptarse para pronosticar series de tiempo económico. Dos de los filtros mejor conocidos son Wiener y Kalman. Recientemente ha surgido interés por las redes neuronales, un método basado en la investigación por computadora. Ninguna de estas técnicas resulta asequible al pronóstico automático, y todas requieren de un usuario sofisticado y conocedor. La simulación de Monte Carlo es otra herramienta útil para el pronóstico en ambientes complejos. Este capítulo concluye con una discusión sobre varias cuestiones en torno a los pronósticos, que surgen en el contexto del control de inventarios.

Como Charles F. Kettering dijo de manera tan elocuente, “Mi preocupación radica en el futuro, ya que planeo pasar el resto de mi vida ahí”. Pero el futuro nunca puede conocerse, por eso lo predecimos. En concordancia con esto, pronosticamos patrones de tráfico y planes de ruta.

Pronosticamos cuáles serán los mejores alimentos en un restaurante en particular, y ordenamos de acuerdo con ello. Elegimos las universidades a las que nos gustaría asistir según el pronóstico basado en nuestras experiencias ahí y las puertas que un título de esa universidad nos puede abrir. Realizamos cientos de pronósticos todos los días, algunos pensados con mucho cuidado, algunos hechos más inconscientemente. El pronóstico juega un papel central en nuestras vidas.

Del mismo modo, el pronóstico juega un papel central en la función de operaciones de una compañía. Toda la planeación de negocios se basa en pronósticos. Las ventas de productos nuevos y existentes, los requerimientos y disponibilidades de materias primas, las habilidades cambiantes de los trabajadores, las necesidades de capacidad y la política internacional son sólo algunos de los factores que pueden afectar el éxito futuro de la compañía.

Las dos áreas funcionales de la compañía que mayor uso hacen de los pronósticos son mercadotecnia y producción. Mercadotecnia normalmente pronostica las ventas de líneas de producto nuevas y existentes. El departamento de producción usa los pronósticos de ventas para la planeación de operaciones. En algunos casos los pronósticos preparados por mercadotecnia pueden no cumplir con las necesidades de producción. Por ejemplo, para determinar los niveles adecuados de inventario de refacciones, es necesario conocer los calendarios de mantenimiento y pronosticar las descomposturas de las máquinas. Esta área también puede pronosticar para grupos o familias de productos, aunque la producción requiera pronósticos para unidades individuales de artículos en almacén para su planeación.

Se ha visto cómo ciertas compañías se benefician de un buen pronóstico, mientras otras deben pagar las consecuencias de uno malo. Durante la década de 1960, el gusto de los consumidores por los automóviles cambió poco a poco: de grandes y pesados devoradores de combustible a vehículos más pequeños y con mejor aprovechamiento de la gasolina. Detroit se tardó en responder a este cambio y sufrió el embargo petrolero de la OPEC que repercutió a fines de la década siguiente, cuando el gusto dio un dramático giro hacia autos más pequeños. Compaq Computer se convirtió en líder del mercado a principios de los ochenta al pronosticar adecuadamente que los consumidores demandarían una versión portátil de la PC de IBM, lo que dio a Compaq una popularidad más allá de lo esperado. El pronóstico jugó cierto papel en el éxito inicial de Ford Motors y en su posterior pérdida de vigencia. Henry Ford vio que el consumidor deseaba un auto más sencillo y menos costoso, más fácil de mantener que los que ofrecían la mayor parte de los fabricantes a principios del siglo pasado. Su Modelo T dominó la industria. Sin embargo, no se dio cuenta de que los consumidores se cansarían del diseño abierto del Modelo T. Ford no pronosticó que el consumidor desearía otros diseños, lo que casi provocó el fin de la compañía que había monopolizado la industria sólo unos cuantos años atrás.

¿Pueden pronosticarse con exactitud todos los eventos? Ciertamente, la respuesta es no. Consideremos el experimento de lanzar una moneda al aire. Suponiendo que se trata de una moneda en buen estado y que el acto de lanzarla al aire no introduce ningún sesgo, lo mejor que puede pensarse es que existen las mismas posibilidades de que salga cualquiera de los lados de la moneda. Nadie ha sido capaz de superar en forma consistente el porcentaje de predicción de 50% en el caso de un experimento como este durante un tiempo demasiado largo. Muchos fenómenos reales se describen de manera muy exacta por medio de este tipo de experimento. Los juegos de posibilidades que pueden encontrarse en los casinos son aleatorios. Al poner las probabilidades a su favor, la casa siempre garantiza su triunfo a largo plazo. Existe evidencia de que los precios diarios de las acciones siguen un proceso puramente aleatorio, como el experimento de la moneda. Los estudios han demostrado que los manejadores profesionales de dinero en raras ocasiones funcionan mejor que los portafolios que se generan de forma puramente aleatoria.

En la administración de operaciones y de la producción, en principio importa el pronóstico de la demanda de producto. Debido a que la demanda tiene muchas posibilidades de ser aleatoria bajo muchas circunstancias, ¿pueden los métodos de pronóstico proporcionar cualquier valor? En la mayoría de los casos, la respuesta es sí. Aunque algunas porciones del proceso de demanda pueden ser impredecibles, es posible predecir otras partes. Si se contemplan las tendencias, los ciclos y la variación estacional tendremos una ventaja para prevenir el resultado del "lanzamiento de la moneda". En este capítulo consideramos métodos para predecir valores futuros de una serie observando cómo ha sido el pasado.

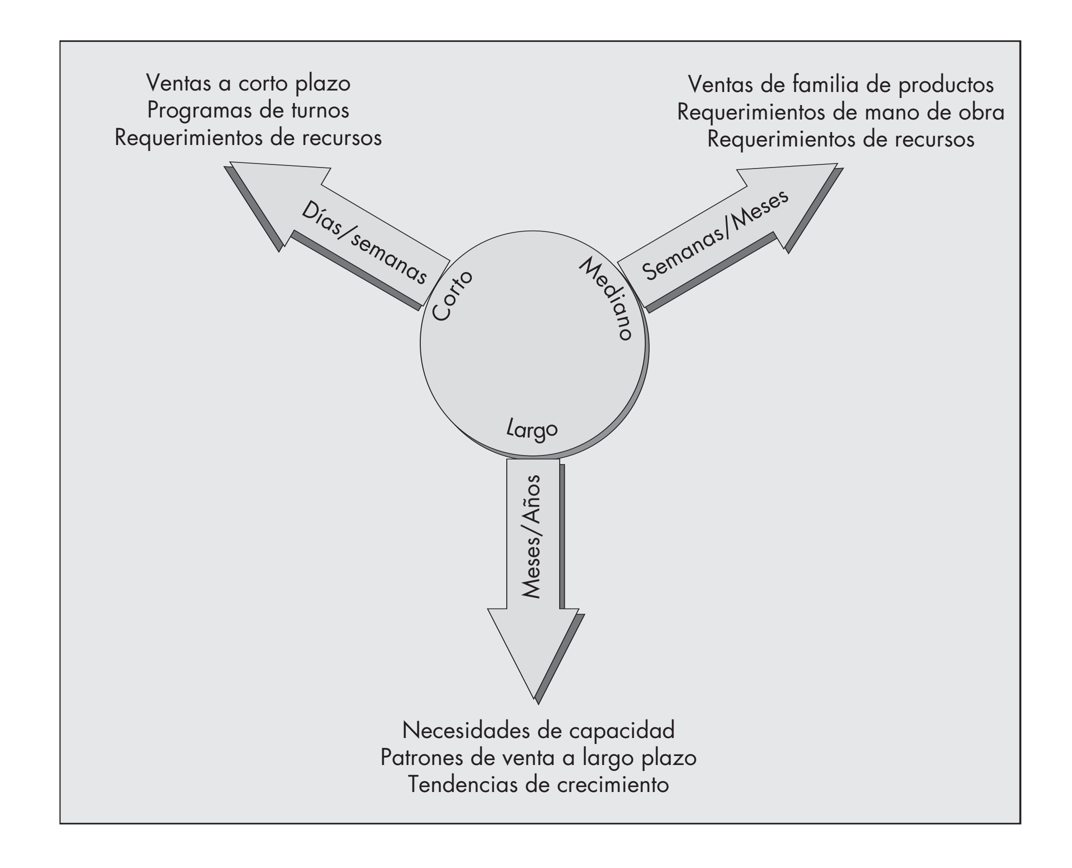
2.1 EL HORIZONTE DE TIEMPO EN LOS PRONÓSTICOS

Podemos clasificar los problemas de pronóstico de acuerdo con varias dimensiones. Una es el horizonte de tiempo. En la figura 2-1 se presenta un esquema que muestra los tres horizontes cronológicos relacionados con el pronóstico y los problemas normales del pronóstico que se encuentran en la planeación de operaciones asociadas con cada uno. Los pronósticos a corto plazo son cruciales para la planeación día con día, y normalmente son medidos en días o semanas. Son de utilidad para la administración de inventarios; para planes de producción que pueden derivarse de un sistema de planeación de requerimientos de materiales (que se comentará con mayor detalle en el capítulo 7); y para la planeación de requerimientos de recursos. La programación de turnos puede requerir que se pronostiquen las preferencias y disponibilidades de los trabajadores.

El mediano plazo se mide en semanas y meses. Los patrones de ventas para las disponibilidades y requerimientos de trabajadores y familias de productos son problemas de pronóstico a mediano plazo que se presentan en la administración de operaciones.

La producción a largo plazo y las decisiones de fabricación, analizadas en el capítulo 1, son parte de la estrategia global de fabricación de la compañía. Un ejemplo es planear a largo plazo las necesidades de capacidad. Cuando se espera que las demandas se incrementen, la compañía debe planificar la construcción de nuevas instalaciones y/o el retroajuste de las instalaciones existentes con nuevas tecnologías. Las decisiones de planeación de capacidad pueden requerir del despido de personal en algunas circunstancias. Por ejemplo, General Motors Corporation históricamente dominó cerca de 45% del mercado interno de automóviles. Sin embargo, en la década de los noventa, el porcentaje se derrumbó al 35%. Como resultado de esto, GM fue forzado a reducir significativamente sus operaciones de fabricación para mantenerse rentable.

FIGURA 2-1
Horizontes de pronóstico en la planeación de operaciones.



2.2 CARACTERÍSTICAS DE LOS PRONÓSTICOS

1. *Normalmente están equivocados*. Por extraño que parezca, esta es probablemente la característica más ignorada y significativa de casi todos los métodos de pronóstico. Los pronósticos, una vez determinados, se consideran información conocida. Los requerimientos de recursos y los calendarios de producción pueden necesitar modificaciones si la demanda de pronóstico prueba ser inexacta. El sistema de planeación debe ser lo suficientemente sólido para ser capaz de reaccionar ante errores de pronóstico no anticipados.

2. *Un buen pronóstico es más que un simple número*. Dado que los pronósticos generalmente están equivocados, un buen pronóstico incluye cierta medida de error. Esto puede ser mediante un rango, o una medición de error como sería la variación de la distribución del error de pronóstico.

3. *Los pronósticos agregados son más exactos*. Recuerde, a partir de las estadísticas, que la variación del promedio de una colección de variables aleatorias independientes distribuidas de manera idéntica es menor que la variación de cada una de las variables aleatorias; esto es, la variación de la muestra media es menor que la variación de población. Este mismo fenómeno se presenta en los pronósticos. En una base porcentual, el error cometido en las ventas pronosticadas para una línea completa de productos generalmente es menor que el error cometido en el pronóstico de ventas para un artículo individual.

4. *Entre más lejano sea el horizonte de pronóstico*, *menos exacta será la predicción*. Esta característica es muy intuitiva. Puede predecirse el valor que tendrá mañana el Dow Jones Industrial Average con mayor exactitud que el del próximo año.

5. *Los pronósticos no deben usarse para excluir información conocida*. Una técnica particular puede resultar en pronósticos razonablemente exactos en la mayoría de las ocasiones. Sin embargo, puede haber información disponible concerniente a la demanda futura que no se mostró en la historia pasada de la serie. Por ejemplo, la compañía puede estar planeando una venta promocional especial para un artículo en particular, por lo que la demanda probablemente será mayor que la normal. Esta información debe incluirse manualmente en el pronóstico.

2.3 MÉTODOS SUBJETIVOS DE PRONÓSTICO

Clasificamos los métodos de pronóstico como **subjetivos** u **objetivos**. Un método de pronóstico subjetivo se basa en el juicio humano. Existen varias técnicas para solicitar opiniones y con base en éstas poder pronosticar:

1. *Agregados de la fuerza de ventas*. Al pronosticar la demanda de un producto, una buena fuente de información subjetiva es la fuerza de ventas de la compañía. La fuerza de ventas tiene contacto directo con los consumidores y, por lo tanto, está en muy buena posición para observar los cambios en sus preferencias. Para desarrollar un pronóstico con la fuerza de ventas, los miembros de ésta presentan sus estimados de ventas de los productos que venderán durante el próximo año. Estos estimados pueden ser números individuales o varios números, como estimados pesimistas, más probablemente, y también optimistas. Los gerentes de ventas deben entonces agregar cálculos individuales para llegar a pronósticos generales en torno a cada región geográfica o grupo de productos. El agregado de fuerza de ventas puede ser inexacto cuando la compensación al personal de ventas se basa en cubrir una cuota. En ese caso, existe un claro incentivo para que la fuerza de ventas realice deliberadamente una estimación a la baja.

2. *Encuestas al cliente*. Éstas pueden señalar tendencias futuras y cambiar los patrones de preferencias. Sin embargo, para que las encuestas y los planes de muestreo sean efectivos, deben diseñarse con mucho cuidado para garantizar que los datos resultantes sean estadísticamente

imparciales y representativos de la base de clientes. Los cuestionarios mal diseñados o un esquema de muestreo no válido pueden provocar conclusiones erróneas.

3. *Juicio de opinión ejecutiva*. Cuando no existe registro del pasado, como en el caso de nuevos productos, la opinión de un experto puede ser la única fuente de información para preparar pronósticos. El enfoque aquí es combinar sistemáticamente las opiniones de los expertos para obtener un pronóstico. Para la planeación de nuevos productos, es necesario solicitar las opiniones del personal en las áreas de mercadotecnia, finanzas y producción. La combinación de pronósticos individuales puede hacerse de varias formas. Una es realizar las entrevistas de pronóstico directamente a los ejecutivos y desarrollar un pronóstico a partir de los resultados que se obtengan de éstas. Otra es solicitar a los ejecutivos que se reúnan como grupo y lleguen a un consenso.

4. *El método Delphi*. El método Delphi, al igual que el método del juicio de opinión ejecutiva, se basa en solicitar las opiniones de los expertos. La diferencia radica en la forma en la que se combinan las opiniones individuales. (El método se denomina así por el oráculo de Delfos de la antigua Grecia, que supuestamente tenía el poder de predecir el futuro.) El método Delphi intenta eliminar algunas carencias inherentes a la dinámica de grupos, en la que las personalidades de algunos miembros del grupo dominan las de otros miembros. Se solicita a un grupo de expertos que exprese sus opiniones, de preferencia por medio de una encuesta de muestra individual. Las opiniones se recopilan y entonces se devuelve un resumen de los resultados a los expertos, con especial atención a aquellas opiniones que son significativamente diferentes de los promedios de grupo. Se pregunta a los expertos si desean reconsiderar sus opiniones originales en virtud de la respuesta obtenida. El proceso se repite hasta que (idealmente) se alcanza un consenso general.

Al igual que con cualquier otra técnica en particular, el método Delphi tiene ventajas y desventajas. Su principal ventaja es que es un medio para evaluar la opinión individual sin las preocupaciones normales de las interacciones personales. En cuanto a su lado negativo, el método es altamente sensible al cuidado en la formulación del cuestionario. Debido a que las discusiones se han excluido intencionalmente del proceso, los expertos no cuentan con ningún mecanismo para resolver aspectos ambiguos. Aún más, no es realmente necesario llegar a un consenso de grupo en algún momento. Un interesante caso de estudio sobre una aplicación exitosa del método Delphi se puede encontrar en Helmer y Rescher (1959).

2.4 MÉTODOS OBJETIVOS DE PRONÓSTICO

Los métodos objetivos de pronóstico son aquellos en los que el pronóstico se deriva de un análisis de datos. Un método de **series de tiempo** es aquel que usa sólo valores pasados en cuanto al fenómeno que deseamos predecir. Los **modelos causales** son aquellos que usan datos provenientes de fuentes distintas a las series que están pronosticando: esto es, pueden existir otras variables con valores que están *vinculadas* de alguna forma a lo que se está pronosticando. Primero analizaremos ese tipo de métodos.

Modelos causales

Sea Y el fenómeno que deseamos pronosticar y $X_1, X_2, \ldots, X_n$ son n variables que creemos que están relacionadas con Y. Entonces un modelo causal es aquel en el que el pronóstico para Y es cierta función de estas variables, digamos,

$$Y = f(X_1, X_2, \ldots, X_n).$$

Los modelos econométricos son modelos causales especiales en los que la relación entre Y y $(X_1, X_2, \ldots, X_n)$ es lineal. Esto es,

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n$$

para algunas constantes ($\alpha_1, \ldots, \alpha_n$). El método de mínimos cuadrados es el que se usa más a menudo para encontrar estimadores para las constantes. (El método se analiza en el apéndice 2-B para el caso de una variable independiente).

Consideremos un ejemplo sencillo de un modelo de pronóstico causal. Un corredor de bienes raíces está tratando de calcular su ingreso para el próximo año. En el pasado ha observado que su ingreso es prácticamente proporcional al número total de ventas de viviendas en su territorio. También se ha dado cuenta de que normalmente ha habido una relación cercana entre las ventas de vivienda y las tasas de interés para las hipotecas a casas. Puede construir un modelo de la forma

$$Y_t = \alpha_0 + \alpha_1 X_{t-1},$$

donde Y_t es el número de ventas en el año t y X_{t-1} es la tasa de interés en el año $t - 1$. Con base en datos pasados podrá determinar entonces los estimadores de mínimos cuadrados para las constantes α_0 y α_1. Supongamos que los valores de estos estimadores son actualmente $\alpha_0 = 385.7$ y $\alpha_1 = -1\,878$. Entonces, la relación estimada entre las ventas de casas y las tasas de interés hipotecario es

$$Y_t = 385.7 - 1\,878 X_{t-1},$$

donde X_{t-1}, la tasa de interés del año anterior, se expresa como un decimal. Entonces, si la tasa de interés actual es de 10%, el modelo pronosticará que el número de ventas para el siguiente año en su territorio será de $385.7 - 187.8 = 197.9$, o cerca de 198 casas vendidas.

Los modelos causales de este tipo son comunes para predecir fenómenos económicos como el producto nacional bruto (PNB) y el producto interno bruto (PIB). Tanto el MIT como la Wharton School of Business de la Universidad de Pennsylvania han desarrollado modelos econométricos a gran escala para hacer estas predicciones. Estos modelos normalmente se utilizan en las áreas de finanzas y economía de la compañía para pronosticar valores de variables macroeconómicas. Los métodos de series de tiempo se utilizan comúnmente en aplicaciones de planeación de operaciones.

Métodos de series de tiempo

Con frecuencia, los métodos de series de tiempo se califican de ingenuos, puesto que no requieren más información que los valores pasados de la variable que se va a predecir. *Series de tiempo* es simplemente un término elegante para hacer referencia a un conjunto de fenómenos físicos o económicos observados en puntos discretos de tiempo, normalmente espaciados equitativamente. La idea es que la información del patrón de observaciones pasadas puede inferirse y usarse para pronosticar valores futuros de las series.

En el análisis de series de tiempo intentamos aislar los patrones que surgen con mayor frecuencia. Éstos incluyen los siguientes:

1. *Tendencia*. Se refiere a la proclividad de una serie de tiempo a exhibir un patrón estable de crecimiento o de declive. Distinguimos entre la tendencia lineal (el patrón descrito por una línea recta) y la tendencia no lineal (el patrón descrito por una función no lineal, como una curva exponencial o cuadrática). Cuando no se especifica el patrón de la tendencia, generalmente se da por hecho que es lineal.

2. *Estacionalidad*. Un patrón estacional es aquel que se repite en intervalos fijos. En las series de tiempo, generalmente pensamos en el patrón que se repite cada año, aunque también son comunes los patrones estacionales mensuales, semanales y diarios. La moda, el helado y el combustible muestran un patrón estacional anual. El consumo de electricidad muestra un fuerte patrón estacional diario.

Aplicación panorámica

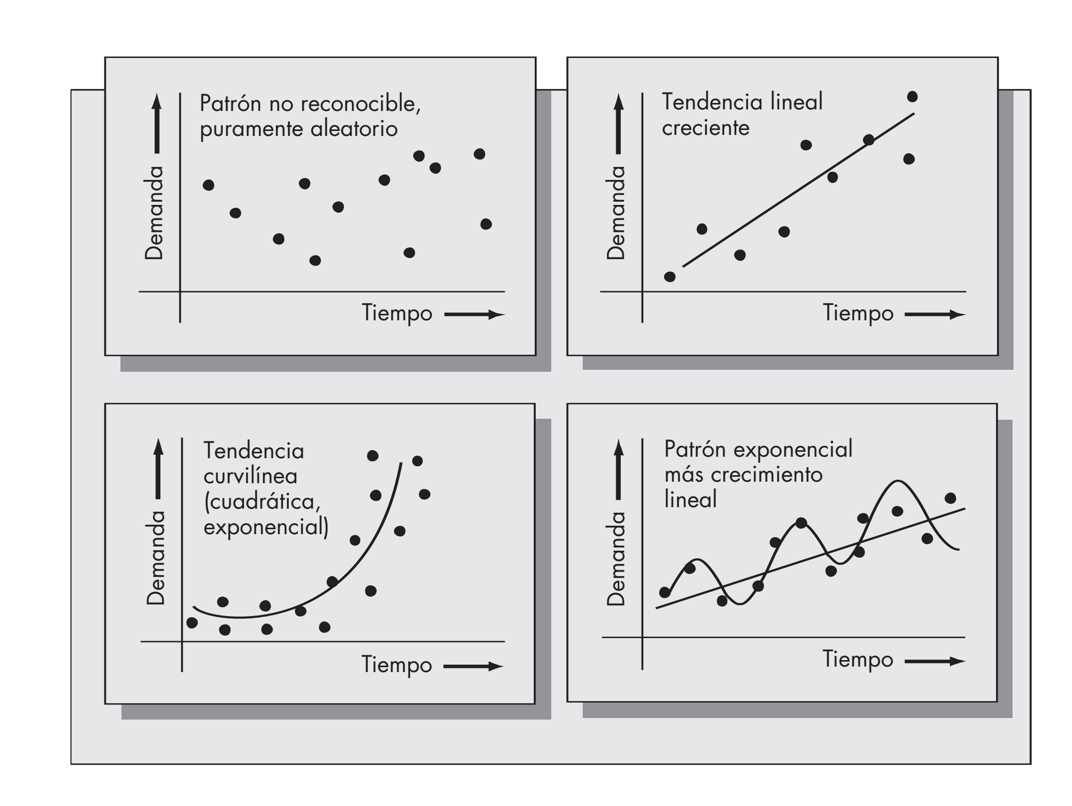
ADVANCED FORECASTING INC. DA SERVICIO A LA INDUSTRIA DE LOS SEMICONDUCTORES

Advanced Forecasting, Inc. (AFI), es una compañía con sede en Cupertino, que se especializa en proporcionar pronósticos para las ventas de semiconductores e industrias afines, como proveedores y equipo de semiconductores. La compañía cuenta con un historial de aciertos en la predicción de momentos decisivos en los patrones de ventas de semiconductores durante más de una década. Los pronósticos están determinados por modelos cuantitativos (como los analizados en este capítulo). Aunque la empresa se reserva los modelos reales utilizados, los pronósticos se apoyan en factores económicos básicos relacionados con la industria de los semiconductores. De acuerdo con el fundador de la compañía, el Dr. Moshe Handelsman, el problema con la mayor parte de las predicciones de los pronosticadores es que se basan en opiniones subjetivas y datos cualitativos. AFI usa un modelo matemático para obtener sus pronósticos, que no son producto de la adivinación.

Aunque la firma es sólo un pequeño jugador en la arena del pronóstico, su éxito ha sido impresionante. Han sido capaces de predecir consistentemente los principales cambios en el mercado de los semiconductores, lo cual es fundamental para la administración. De acuerdo con Jean-Philippe Dauvin, vicepresidente y economista en jefe de SGS-Thomson Microelectronics: "Nuestra dirección general presta más atención a las predicciones de Advanced Forecasting que a cualquier otra fuente de la industria". Los pronósticos exactos permiten a la administración enfrentar importantes cuestiones estratégicas como el momento de expandir la capacidad de producción, las necesidades de personal, y las demandas en mercadotecnia y ventas. El éxito de AFI demuestra que el pronóstico basado en aspectos cuantitativos puede proporcionar pronósticos exactos de forma consistente y que, a largo plazo, son mucho más confiables que los métodos subjetivos.

Fuente: Advanced Forecasting, Inc., sitio web, http://www.*adv-forecast.com*/afi/

FIGURA 2-2
Patrones de series de tiempo.



3. *Ciclos*. La variación cíclica es similar a la estacionalidad, excepto porque la duración y la magnitud del ciclo pueden variar. Los ciclos se asocian con variaciones económicas a largo plazo (esto es, los ciclos comerciales) que pueden presentarse además de las fluctuaciones estacionales.

4. *Aleatoriedad*. Una serie aleatoria pura es aquella en la que no existe un patrón reconocible para los datos. Los datos pueden generarse de una forma que, aun siendo puramente aleatoria, muchas veces aparenta tener una estructura. Un ejemplo sería la metodología de tabuladores del mercado de valores que impone formas de patrones aleatorios en los datos de precios del mercado. Del otro lado de la moneda, los datos que parecen ser aleatorios pueden tener una estructura definitiva. Los datos verdaderamente aleatorios que fluctúan en torno a una media fija forman lo que se conoce como patrón horizontal.

En la figura 2-2 se presenta una serie de tiempo en la que se muestran estos patrones.

Problemas para las secciones 2.1 a 2.4

1. Mencione los cuatro componentes de las series de tiempo (es decir, los cuatro patrones distintos exhibidos por las series de tiempo).
2. ¿Cuál es la diferencia entre la estacionalidad y los ciclos en el análisis de series de tiempo?
3. ¿Cuál es el tipo apropiado de método de pronóstico que debe usarse en cada uno de los siguientes escenarios?

   a) Holiday Inn, Inc., está intentando predecir la demanda que tendrán el próximo año las habitaciones de los moteles, tomando como base un historial de observaciones de demanda.
   b) Standard Brands ha desarrollado un nuevo tipo de pintura exterior. La compañía desea pronosticar las ventas con base en los nuevos proyectos de construcción de viviendas.
   c) IBM está tratando de averiguar el costo del faltante de existencias de un componente crítico para las unidades de cinta. Esto, por medio de una encuesta a gerentes en diferentes centros de refacciones a nivel nacional. Las encuestas se envían de regreso a los gerentes para una nueva valoración y se repite el proceso hasta que se alcanza un consenso.

4. Comente el papel de los pronósticos en las siguientes funciones de la compañía:

   a) Mercadotecnia
   b) Contabilidad
   c) Finanzas
   d) Producción

5. Distinga entre los siguientes tipos de pronósticos:

   a) Artículo único contra agregados.
   b) A corto plazo contra largo plazo.
   c) Causal contra ingenuo.

6. ¿Cuál es la ventaja del método Delphi sobre el método del juicio de opinión ejecutiva? ¿Qué tienen en común estos métodos?
7. Consideremos el problema de elegir una universidad adecuada cuando se es alumno de preparatoria. ¿Cuáles serían sus preocupaciones respecto al pronóstico al momento de tomar esa decisión? En particular, enumere los pronósticos a corto, mediano y largo plazo que puede haber considerado al tomar su decisión final. ¿Qué fuentes objetivas de datos habría utilizado para obtener mejores pronósticos en cada caso?
8. Comente la siguiente frase expresada por un gerente de control de inventarios: "No es mi culpa que se nos hayan acabado estas refacciones. El pronóstico de la demanda estaba equivocado".
9. Comente la siguiente frase: "Los economistas predicen que las tasas de interés continuarán por debajo de 10% cuando menos durante quince años más".

2.5 CONVENCIONES DE NOTACIÓN

El siguiente análisis se refiere a los métodos de series de tiempo. Defina $D_1, D_2, \ldots, D_t, \ldots$ como los valores de demanda observados durante los periodos $1, 2, \ldots, t, \ldots$. Asumiremos en todo momento que $\{D_t, t \geq 1\}$ es la serie de tiempo que se pretende predecir. Aún más, supondremos que si estamos realizando un pronóstico en el periodo t, entonces hemos observado $D_t, D_{t-1}, \ldots$ pero no hemos observado D_{t+1}.

Defina F_t como el pronóstico hecho *para* el periodo t en el periodo $t - 1$. Es el pronóstico hecho al final del periodo $t - 1$ después de haber observado $D_{t-1}, D_{t-2}, \ldots$, pero antes de observar D_t. Por lo tanto, estamos suponiendo que todos los pronósticos son pronósticos de un paso adelante; esto es, que están hechos para la demanda en el próximo periodo. Los pronósticos de paso adelante múltiple serán comentados en una sección posterior.

Finalmente, observemos que se obtiene un pronóstico de serie al aplicar un cierto conjunto de pesos a los datos pasados. Esto es,

$$F_t = \sum_{n=0}^{\infty} a_n D_{t-n} \qquad \text{para algún conjunto de pesos } a_0, a_1, \ldots.$$

Los métodos de series de tiempo analizados en este capítulo se distinguen sólo por la elección de los pesos.

2.6 EVALUACIÓN DE PRONÓSTICOS

Defina el error de pronóstico en cualquier periodo t, e_t, como la diferencia entre el pronóstico para el periodo y la demanda real para el mismo periodo. Para pronósticos de varios pasos adelante,

$$e_t = F_{t-\tau,\, t} - D_t,$$

y para los pronósticos de un solo paso adelante

$$e_t = F_t - D_t.$$

Digamos que $e_1, e_2, \ldots, e_n$ son los errores de pronóstico observados en n periodos. Las dos medidas más comunes de exactitud de pronóstico durante estos periodos n son la desviación absoluta media (DAM) y el error cuadrático medio (ECM), dados por las fórmulas siguientes:

$$\text{DAM} = (1/n) \sum_{i=1}^{n} |e_i|$$

$$\text{ECM} = (1/n) \sum_{i=1}^{n} e_i^2$$

Observe que el ECM es similar a la varianza de una muestra aleatoria. La DAM es con frecuencia el método preferido para medir el error de pronóstico debido a que no requiere la elevación al cuadrado. Es más, cuando los errores de pronóstico se distribuyen normalmente, como generalmente se supone, un estimado de la desviación estándar del error de pronóstico, σ_e, está dado por 1.25 veces la DAM.

Aunque la DAM y el ECM son las dos medidas más comunes de la exactitud del pronóstico, también se utilizan otras medidas. Una que no es dependiente de la magnitud de los valores de demanda se conoce como el error porcentual absoluto medio (EPAM) y está dado por la fórmula

$$\text{EPAM} = \left[(1/n) \sum_{i=1}^{n} |e_i / D_i|\right] \times 100.$$

Ejemplo *2.1*

Artel, un fabricante de memorias de acceso aleatorio estático (SRAM), cuenta con plantas de producción en Austin, Texas, y Sacramento, California. Los gerentes de esas plantas tienen que pronosticar los rendimientos de producción (medidos en porcentaje) para sus plantas con una semana de antelación. Basada en seis pronósticos semanales, la dirección de la compañía desea determinar qué gerente tiene más éxito para predecir los rendimientos de su planta. Los resultados de sus predicciones se muestran en la tabla siguiente:

Semana	P1	O1	|E1|	|E1/O1|	P2	O2	|E2|	|E2/O2|
1	92	88	4	.0455	96	91	5	.0549
2	87	88	1	.0114	89	89	0	.0000
3	95	97	2	.0206	92	90	2	.0222
4	90	83	7	.0843	93	90	3	.0333
5	88	91	3	.0330	90	86	4	.0465
6	93	93	0	.0000	85	89	4	.0449

Interprete P1 como el pronóstico hecho por el gerente de la planta 1 al principio de cada semana, O1 como el rendimiento observado al final de cada semana en la planta 1, y E1 como la diferencia entre los rendimientos observados y los pronosticados. Las mismas definiciones se aplican a la planta 2.

Comparemos el desempeño de estos gerentes usando las tres medidas, DAM, ECM y EPAM, tal como se definieron anteriormente. Para calcular la DAM simplemente promediamos los errores absolutos observados:

$$\text{DAM}_1 = 17/6 = 2.83$$

$$\text{DAM}_2 = 18/6 = 3.00.$$

Con base en las DAM, el primer gerente tuvo una ligera ventaja. Para calcular el ECM en cada caso, eleve al cuadrado los errores observados y promedie los resultados para obtener

$$\text{ECM}_1 = 79/6 = 13.17$$

$$\text{ECM}_2 = 70/6 = 11.67.$$

Los pronósticos del segundo gerente tuvieron un ECM menor que el primero, aun cuando las DAM fueron hacia el otro lado. ¿Por qué el cambio? La razón por la que ahora el primer gerente pronostica peor reside en que el ECM es más sensible a un error grande de lo que lo es la DAM. Observe que 7, el error más grande que se observó, correspondió al gerente 1.

Comparemos ahora sus desempeños con base en el EPAM. Para calcular el EPAM, promediamos los porcentajes de errores y los rendimientos observados:

$$\text{EPAM}_1 = .0325$$

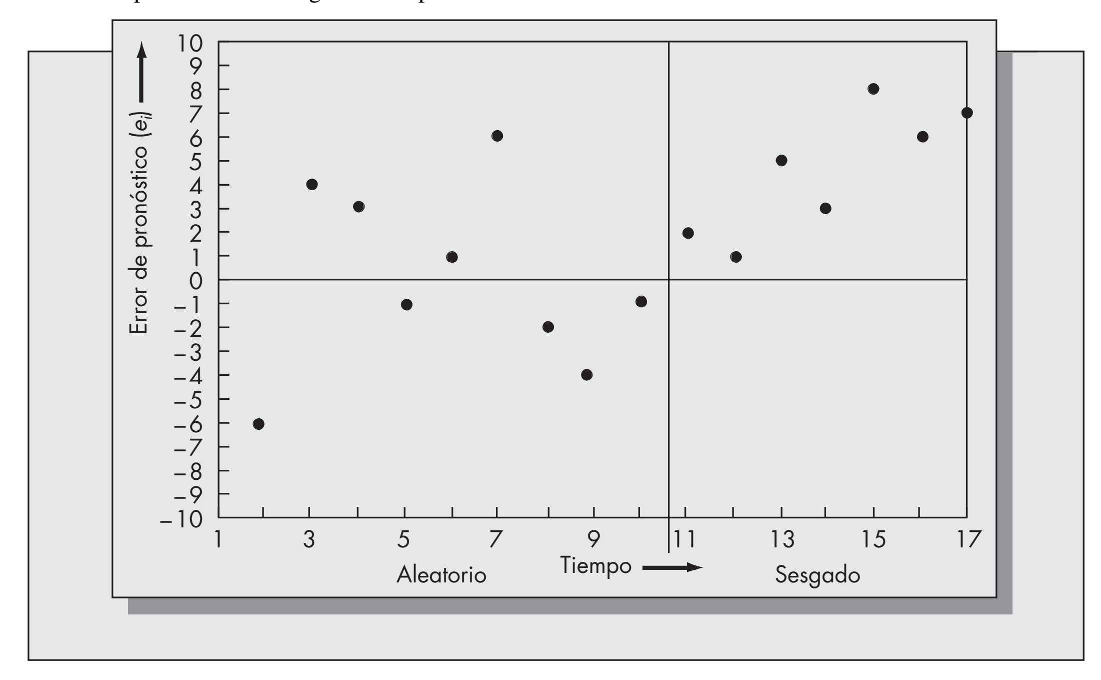
$$\text{EPAM}_2 = .0336.$$

Al utilizar la DAM o el EPAM, el primer gerente tiene una pequeña ventaja, pero cuando usa el ECM, el segundo gerente logra mejores resultados. Las habilidades de pronóstico de los dos gerentes parecen ser muy similares. Quién debe ser declarado "vencedor" depende de cuál método de evaluación elija la dirección de la empresa.

Una propiedad deseable para los pronósticos es que sean insesgados. Matemáticamente, eso significa que $E(e_i) = 0$. Una forma de dar seguimiento a un método de pronóstico es graficar los valores del error de pronóstico e_i a lo largo del tiempo. Si el método es insesgado, los errores de pronóstico deben fluctuar aleatoriamente por encima y por debajo de cero. En la figura 2-3 se presenta un ejemplo.

Una alternativa al método gráfico es calcular la suma acumulada de los errores de pronóstico, Σe_i. Si el valor de esta suma se desvía sustancialmente de cero (ya sea hacia arriba o hacia abajo), es un indicativo de que el método de pronóstico está sesgado. Al final de este capítulo analizamos una técnica de suavizamiento que puede usarse también para señalar un sesgo en los pronósticos. Las tablas de control estadístico se utilizan para identificar valores inusualmente grandes del error de pronóstico. (Las tablas de control estadístico se comentan en el capítulo 11).

FIGURA 2-3
Errores de pronóstico a lo largo del tiempo.



Problemas para la sección 2.6

10. En *Dave Pelz's Short Game Bible*, el autor intenta caracterizar la habilidad de un golfista desde una distancia específica en términos de la proporción de error en el tiro y la distancia del intento. Por ejemplo, un hierro cinco que se golpea a 175 yardas y queda 20 yardas fuera del objetivo tiene una proporción de exactitud de 20/175 = .114, mientras que un tiro a 60 yardas desde la arena que queda a 10 yardas del objetivo tendrá una proporción de exactitud de 10/60 = .1667 (entre más baja es la calificación, mejor). ¿A cuál de los métodos de evaluación analizados en esta sección se asemeja más? ¿Por qué este método de evaluación tiene más sentido en el golf que los errores cuadráticos o absolutos?

11. Un método de pronóstico utilizado para predecir las ventas de un abrelatas aplica el siguiente conjunto de pesos a los últimos cinco periodos de datos: .1, .1, .2, .2, .4 (donde .4 se aplica a la observación más reciente). Los valores observados del abrelatas son

Periodo:	1	2	3	4	5	6	7	8
Observación:	23	28	33	26	21	38	32	41

Determine lo siguiente:

a) El pronóstico de un solo paso adelante para el periodo 9.

b) El pronóstico de un solo paso adelante para el periodo 6.

12. Un método de pronóstico simple para las ventas semanales de unidades de disco flexible usado por un proveedor local consiste en obtener el promedio de las dos cifras de ventas más recientes.

Supongamos que las ventas de las unidades de disco durante las 12 semanas anteriores fueron

Semana:	1	2	3	4	5	6	7	8	9	10	11	12
Ventas:	86	75	72	83	132	65	110	90	67	92	98	73

a) Determine los pronósticos de un paso adelante realizados para los periodos 3 a 12 utilizando este método.

b) Determine los errores de pronóstico para estos periodos.

c) Calcule la DAM, el ECM y el EPAM con base en los errores de pronóstico determinados en el inciso *b*.

13. Se han utilizado dos métodos de pronóstico para evaluar la misma serie económica de tiempo. Los resultados son

Pronóstico con el método 1	Pronóstico con el método 2	Valor realizado de la serie
223	210	256
289	320	340
430	390	375
134	112	110
190	150	225
550	490	525

Compare la efectividad de estos métodos mediante el cálculo del ECM, la DAM y el EPAM. ¿Indican cada una de esas medidas de exactitud del pronóstico que la misma técnica es la mejor? Si no es así, ¿por qué?

14. ¿Qué significa el término *sesgado* en referencia con una técnica de pronóstico en particular?

15. ¿Cuál es el estimado de la desviación estándar del error de pronóstico obtenido con los datos del problema 12?

2.7 MÉTODOS PARA PRONOSTICAR SERIES ESTACIONARIAS

En esta sección analizaremos dos técnicas populares, los promedios móviles y el suavizamiento exponencial para pronosticar series de tiempo exponencial. Una serie de tiempo exponencial es aquella en la que cada observación puede representarse por medio de una constante más una fluctuación aleatoria. En símbolos,

$$D_t = \mu + \epsilon_t,$$

donde μ es una constante desconocida que corresponde a la media de la serie y ϵ_t es un error aleatorio con media cero y varianza σ^2.

Los métodos que consideramos en esta sección se conocen más precisamente como suavizamiento simple o sencillo, y promedios móviles simples o sencillos. Además, los promedios móviles simples incluyen también promedios móviles ponderados, que no describimos aquí. Por conveniencia, no usaremos los modificadores *simple* y *sencillo* en lo que presentamos a continuación. El significado de los términos quedará claro a partir del contexto.

Promedios móviles

Un método de pronóstico sencillo pero popular es el método de los promedios móviles. Un promedio móvil del orden N es sencillamente el promedio aritmético de las observaciones N más recientes.

Para el tiempo dado restringimos la atención a los pronósticos de un paso adelante. Entonces F_t, el pronóstico hecho en el periodo $t - 1$ para el periodo t, está dado por

$$F_t = (1/N) \sum_{i=t-N}^{t-1} D_i = (1/N)(D_{t-1} + D_{t-2} + \cdots + D_{t-N}).$$

En palabras, esto dice que la media de las observaciones N más recientes se utiliza como el pronóstico para el próximo periodo. Usaremos la anotación PM(N) para los promedios móviles del periodo N.

Ejemplo *2.2*

Los datos trimestrales para las fallas de ciertos motores de aeronaves en una base local militar durante los pasados dos años son 200, 250, 175, 186, 225, 285, 305, 190. Los promedios trimestrales y semestrales se usan para pronosticar el número de fallas de motores. Determine los pronósticos de un paso adelante para los periodos 4 a 8 utilizando los promedios móviles de tres periodos, y los pronósticos de un paso adelante para los periodos 7 y 8 utilizando promedios móviles para seis periodos.

Solución

El pronóstico de promedio móvil de tres periodos para el periodo 4 se obtiene mediante el promedio de los tres primeros datos.

$$F_4 = (1/3)(200 + 250 + 175) = 208.$$

El pronóstico de promedio móvil de tres periodos para el periodo 5 es

$$F_5 = (1/3)(250 + 175 + 186) = 204.$$

El pronóstico de promedio móvil de seis periodos para el periodo 7 es

$$F_7 = (1/6)(200 + 250 + 175 + 186 + 225 + 285) = 220.$$

Otros pronósticos se calculan de manera similar. Si disponemos los pronósticos y los errores de pronóstico asociados, obtenemos

Trimestre	Fallas de motor	PM(3)	Error	PM(6)	Error
1	200				
2	250				
3	175				
4	186	208	22		
5	225	204	−21		
6	285	195	−90		
7	305	232	−73	220	−85
8	190	272	82	238	48

Una pregunta interesante es, ¿cómo pueden obtenerse pronósticos de varios pasos adelante? Por ejemplo, supongamos que en el ejemplo 2.2 estamos interesados en utilizar PM(3) en el periodo 3 para realizar un pronóstico del periodo 6. Debido a que el método del promedio móvil se basa en el supuesto de que la serie de demanda es estacionaria, el pronóstico realizado en el periodo 3 para *cualquier* otro periodo futuro será el mismo. Esto es, los pronósticos de un paso adelante y de múltiples pasos adelante son idénticos (aunque el pronóstico de un paso adelante generalmente es más exacto). Por lo tanto, el pronóstico PM(3) realizado en el periodo 3 para el periodo 6 es 208. De hecho, el pronóstico PM(3) realizado en el periodo 3 para cualquier periodo posterior al 3 es también 208.

Una desventaja evidente de la técnica de promedio móvil es que debe recalcularse el promedio de las últimas N observaciones cada vez que surge una nueva observación de demanda. Para N grandes puede resultar tedioso. Sin embargo, el nuevo cálculo de todo el promedio de

N periodos no es necesario en cada periodo, porque

$$F_{t+1} = (1/N) \sum_{i=t-N+1}^{t} D_i = (1/N) \left[D_t + \sum_{i=t-N}^{t-1} D_i - D_{t-N} \right]$$

$$= F_t + (1/N)[D_t - D_{t-N}]$$

Esto significa que para un pronóstico de un paso adelante necesitamos calcular sólo la diferencia entre la demanda más reciente y los *N* periodos de demanda anteriores con el fin de actualizar el pronóstico. No obstante, sigue siendo necesario mantener un registro de todas las *N* observaciones pasadas. ¿Por qué?

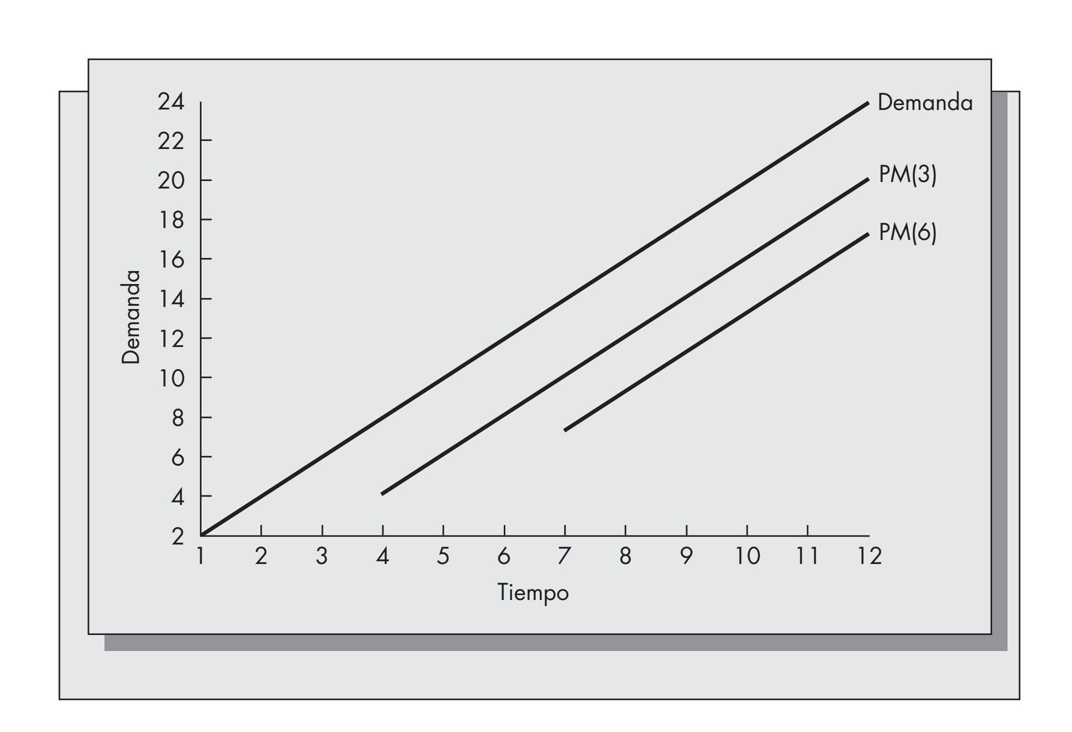
El promedio móvil se queda detrás de la tendencia

Consideremos un proceso de demanda en el que exista una tendencia definitiva. Por ejemplo, supongamos que la demanda observada es 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24. Consideremos los pronósticos PM(3) y PM(6) de un paso adelante para esta serie.

Periodo	Demanda	PM(3)	PM(6)
1	2		
2	4		
3	6		
4	8	4	
5	10	6	
6	12	8	
7	14	10	7
8	16	12	9
9	18	14	11
10	20	16	13
11	22	18	15
12	24	20	17

La demanda y los pronósticos para los periodos respectivos se muestran en la figura 2-4. Obsérvese que los pronósticos PM(3) y PM(6) se quedan detrás de la tendencia. Aún más,

FIGURA 2-4
Los pronósticos de promedio móvil se quedan detrás de la tendencia.



PM(6) tiene un mayor retraso. Esto implica que los promedios móviles simples no son la mejor manera de pronosticar cuando existe una tendencia en las series.

Problemas sobre promedios móviles

Los problemas 16 a 21 se basan en los siguientes datos. Las observaciones de la demanda para una cierta refacción almacenada en el depósito de abastecimiento de refacciones durante el año civil 1999 fueron

Mes	Demanda	Mes	Demanda
Enero	89	Julio	223
Febrero	57	Agosto	286
Marzo	144	Septiembre	212
Abril	221	Octubre	275
Mayo	177	Noviembre	188
Junio	280	Diciembre	312

16. Determine los pronósticos de un paso adelante para la demanda de enero de 2000 utilizando promedios móviles de 3, 6 y 12 meses.
17. Utilizando un promedio móvil de cuatro meses, determine los pronósticos de un paso adelante para el periodo de julio a diciembre de 1999.
18. Utilizando un promedio móvil de cuatro meses, determine los pronósticos de dos pasos adelante para el periodo de julio a diciembre de 1999. (Sugerencia: el promedio de dos pasos adelante para julio se basa en las demandas observadas entre febrero y mayo.)
19. Calcule la DAM para los pronósticos obtenidos en los problemas 17 y 18. ¿Qué método dio mejores resultados? Con base en la teoría de pronósticos, ¿qué método ha dado mejores resultados?
20. Calcule los pronósticos de un paso adelante de promedios móviles de tres y seis meses. ¿Cuál es el efecto que tiene en los pronósticos el incremento de *N* de 3 a 6?
21. ¿Qué significaría un método de pronóstico PM(1)? Compare la exactitud de los pronósticos PM(1) y PM(4) para el periodo de julio a diciembre de 1999.

Suavizamiento exponencial

Otro método de pronóstico muy popular para las series de tiempo estacionarias es el suavizamiento exponencial. El pronóstico actual es el promedio ponderado del último pronóstico y el valor actual de demanda. Esto es,

$$\text{Nuevo pronóstico} = \alpha(\text{Observación actual de demanda}) + (1 - \alpha)\ (\text{Último pronóstico}).$$

En símbolos,

$$F_t = \alpha D_{t-1} + (1 - \alpha)F_{t-1},$$

donde $0 < \alpha \leq 1$ es la constante de suavizamiento, que determina la ponderación relativa colocada en la observación de demanda actual. Interprete $(1 - \alpha)$ como el peso asignado a las observaciones pasadas de la demanda. Por medio de un sencillo reacomodo de términos, la ecuación de suavizamiento exponencial para F_t puede escribirse como

$$\begin{aligned} F_t &= F_{t-1} - \alpha(F_{t-1} - D_{t-1}) \\ &= F_{t-1} - \alpha e_{t-1}. \end{aligned}$$

Escrita de esta forma, vemos que el suavizamiento exponencial puede interpretarse de la siguiente manera: el pronóstico en cualquier periodo t es el pronóstico en el periodo $t - 1$ menos cierta fracción de error de pronóstico observado en el periodo $t - 1$. Nótese que si pronosticamos alto en el periodo $t - 1$, e_{t-1} es positivo y el ajuste consiste en disminuir el pronóstico. De forma similar, si pronosticamos bajo en el periodo $t - 1$, el error es negativo, y el ajuste es incrementar el pronóstico actual.

Como antes, F_t es el pronóstico de un paso adelante para el periodo t realizado en el periodo $t - 1$. Obsérvese que, como

$$F_{t-1} = \alpha D_{t-2} + (1 - \alpha)F_{t-2},$$

esto puede sustituirse en la ecuación anterior para obtener

$$F_t = \alpha D_{t-1} + \alpha(1 - \alpha)D_{t-2} + (1 - \alpha)^2 F_{t-2}.$$

Ahora podemos sustituir F_{t-2} en la misma forma. Si continuamos de esta manera obtenemos la expansión infinita para F_t,

$$F_t = \sum_{i=0}^{\infty} \alpha(1 - \alpha)^i D_{t-i-1} = \sum_{i=0}^{\infty} a_i D_{t-i-1},$$

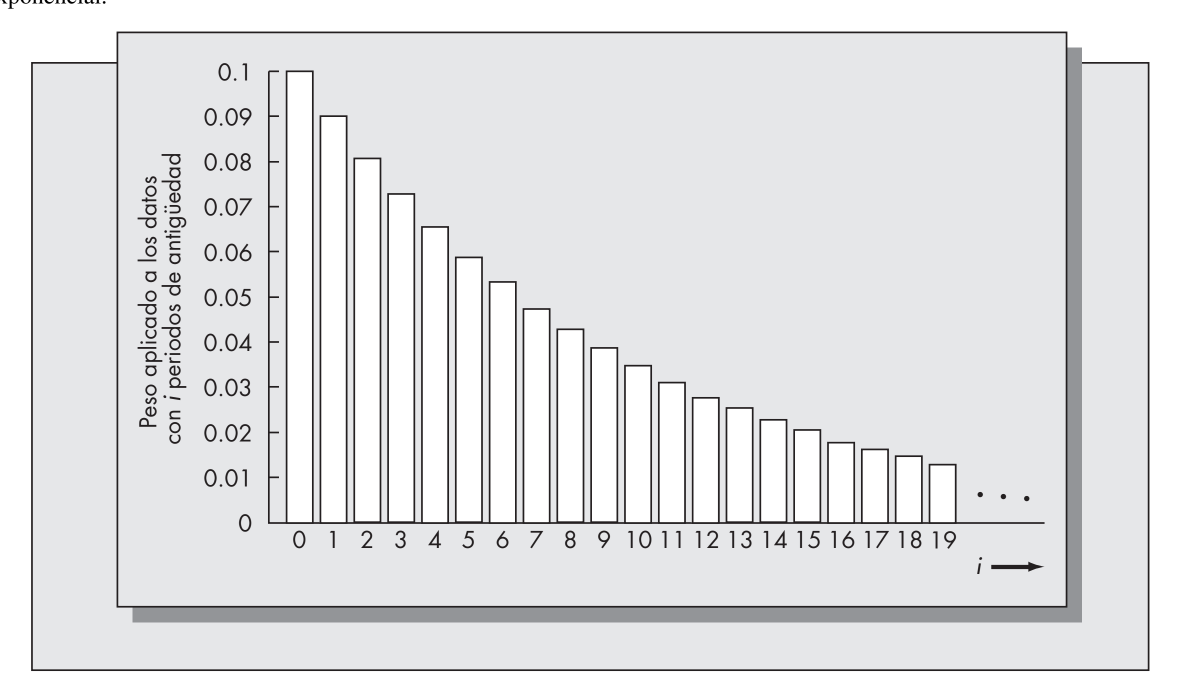
donde los pesos son $a_0 > a_1 > a_2 > \cdots > a_i = \alpha(1 - \alpha)^i$, y

$$\sum_{i=0}^{\infty} a_i = \sum_{i=0}^{\infty} \alpha(1 - \alpha)^i = \alpha \sum_{i=0}^{\infty} (1 - \alpha)^i = \alpha \times 1/[1 - (1 - \alpha)] = 1.$$

Por lo tanto, el suavizamiento exponencial aplica un conjunto de ponderaciones decrecientes a todos los datos pasados. Las ponderaciones se grafican como una función de i en la figura 2-5.

De hecho, podemos ajustar la curva exponencial continua $g(i) = \alpha \exp(-\alpha i)$ a estas ponderaciones, que es la razón por la cual el método se denomina suavizamiento exponencial. La constante de suavizamiento juega esencialmente el mismo papel aquí que el valor de N en los

FIGURA 2-5
Ponderaciones en el suavizamiento exponencial.



promedios móviles. Si α es grande, se realiza mayor ponderación en la observación actual de demanda y menos ponderación sobre las observaciones pasadas, lo que da como resultado pronósticos que reaccionarán rápidamente a los cambios en los patrones de demanda pero pueden tener mucha mayor variación de periodo a periodo. Si α es pequeña, entonces se asigna mayor peso en los datos pasados y los pronósticos son más estables.

Cuando se usa una técnica automática de pronóstico para predecir la demanda en una aplicación de producción, se desea tener pronósticos estables (esto es, pronósticos que no varíen mucho de un periodo a otro). Los pronósticos de demanda se utilizan en el punto inicial para la planeación de producción y programación. La revisión sustancial en estos pronósticos puede causar estragos en los programas de trabajo de los empleados, en las listas de componentes de material y en las órdenes de compra externas. Por esta razón, en general se recomienda un valor de α entre .1 y .2 para las aplicaciones de producción. (Véase por ejemplo, Brown, 1962.)

Los pronósticos de varios pasos adelante se manejan igual para el suavizamiento exponencial que para los promedios móviles; esto es, los promedios de un paso adelante y de varios pasos adelante son iguales.

Ejemplo *2.3*

Consideremos el ejemplo 2.2 (en la página 64), donde los promedios móviles se utilizaron para predecir las fallas de motor de las aeronaves. Los números de fallas observados durante un periodo de dos años fueron 200, 250, 175, 186, 225, 285, 305, 190. Ahora pronosticamos utilizando el suavizamiento exponencial. Con el fin de iniciar el método, asumamos que el pronóstico para el periodo 1 fue 200. Supongamos que $\alpha = .1$. El pronóstico de un paso adelante para el periodo 2 es

$$F_2 = \alpha D_1 + (1 - \alpha)F_1 = (.1)(200) + (.9)(200) = 200.$$

De forma similar,

$$F_3 = \alpha D_2 + (1 - \alpha)F_2 = (.1)(250) + (.9)(200) = 205.$$

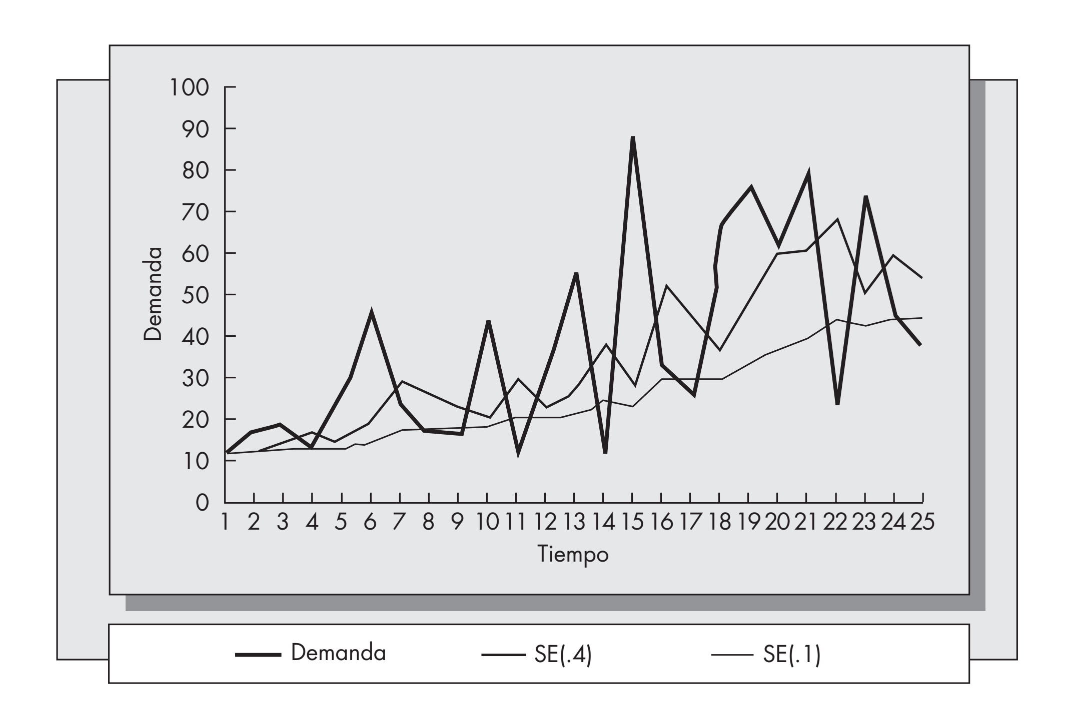
Otros pronósticos de un paso adelante se calculan de la misma manera. Los números de fallas observados y los pronósticos de un paso adelante en cada trimestre son los siguientes:

Trimestre	Fallas	Pronóstico
1	200	200 (por suposición)
2	250	200
3	175	205
4	186	202
5	225	201
6	285	203
7	305	211
8	190	220

Observe el efecto de la constante de suavizamiento. Aunque la serie original muestra una alta varianza, los pronósticos parecen ser muy estables. Repita los cálculos con un valor de $\alpha = .4$. Habrá mucha mayor variación en los pronósticos.

Debido a que el suavizamiento exponencial requiere que en cada etapa tengamos un pronóstico previo, no resulta obvia la forma de iniciar el método. Podemos suponer que el pronóstico inicial es igual al valor inicial de demanda, como lo hicimos en el ejemplo 2.3. Sin embargo, este enfoque tiene un serio inconveniente. El suavizamiento exponencial pone un peso especial en las observaciones pasadas, así que el valor inicial de demanda tendrá un efecto absurdamente grande en los primeros pronósticos. Este problema puede superarse si se per-

FIGURA 2-6
Suavizamiento exponencial para diferentes valores de alfa.



mite que el proceso evolucione durante un número razonable de periodos (digamos 10 o más) y utilizando el promedio aritmético de la demanda durante esos periodos como el pronóstico inicial.

A fin de apreciar el efecto de la constante de suavizamiento, en la figura 2.6 hemos graficado una serie particularmente errática, junto con los pronósticos resultantes, usando valores de $\alpha = .1$ y $\alpha = .4$. Nótese que para $\alpha = .1$ el valor pronosticado de demanda da como resultado un patrón relativamente suave, donde para $\alpha = .4$ el valor pronosticado muestra una variación significativamente mayor. Aunque el suavizamiento con el valor más grande de α hace un mejor trabajo de seguimiento que la serie, para fines de la planeación se persigue la estabilidad proporcionada por una constante de suavizamiento menor.

Ejemplo ***2-3 (continuación)***

Consideremos de nuevo el problema de pronosticar las fallas del motor de aeronaves. Supongamos que estábamos interesados en comparar el desempeño de PM(3) con los pronósticos de suavizamiento exponencial obtenidos anteriormente (SE(.1)). El primer periodo para el cual tenemos un pronóstico usando PM(3) es el periodo 4, así que realizaremos el proceso únicamente para los periodos 4 a 8.

Trimestre	Fallas	PM(3)	|Error|	SE(.1)	|Error|
4	186	208	22	202	16
5	225	204	21	201	24
6	285	195	90	203	82
7	305	232	73	211	94
8	190	272	82	220	30

El promedio aritmético de los errores absolutos, DAM, es 57.6 para el promedio móvil de tres periodos y 49.2 para el suavizamiento exponencial. Los valores respectivos de ECM son 4 215.6 y 3 458.4. Con base únicamente en esta comparación, puede concluirse que el suavizamiento exponencial es un método superior para esta serie. No obstante, esto no es necesariamente cierto.

En el ejemplo 2.3 comparamos el suavizamiento exponencial con $\alpha = .1$ y un promedio móvil con $N = 3$. ¿Cómo sabemos que esa configuración de parámetro es consistente? Los pronósticos PM(3) muestran mucha mayor variabilidad que los pronósticos SE(.1), lo que sugiere que $\alpha = .1$ y $N = 3$ no son valores consistentes de estos parámetros.

La determinación de los valores consistentes de α y N puede realizarse de dos maneras diferentes. Una es equiparar la edad promedio de los datos utilizados para realizar el pronóstico. Un pronóstico con promedios móviles consiste en ponderaciones iguales de $1/N$ aplicadas a las últimas observaciones N. Al multiplicar el peso asignado a cada observación por su "edad", obtenemos la edad promedio de datos para los promedios móviles como

$$\begin{aligned}\text{Edad promedio} &= (1/N)(1 + 2 + 3 + \cdots + N) = (1/N)(N)(N + 1)/2 \\ &= (N + 1)/2.\end{aligned}$$

Para el suavizamiento exponencial, la ponderación aplicada a antiguos periodos de datos i es $\alpha(1 - \alpha)^{i-1}$. Asumamos que tenemos una historia infinita de observaciones pasadas de demanda. Por lo tanto, la edad promedio de datos en un pronóstico de suavizamiento exponencial es

$$\text{Edad promedio} = \sum_{i=1}^{\infty} i\alpha(1 - \alpha)^{i-1} = 1/\alpha.$$

Omitimos los detalles de este cálculo.

Al igualar la edad promedio de datos para los dos métodos, obtenemos

$$\frac{N + 1}{2} = \frac{1}{\alpha},$$

que es equivalente a

$$\alpha = 2/(N + 1) \qquad \text{o} \qquad N = \frac{2 - \alpha}{\alpha}.$$

Por lo tanto, vemos que hubiésemos necesitado un valor de $N = 19$ para $\alpha = .1$ o un valor de $\alpha = .5$ para $N = 3$ para que los métodos fuesen consistentes en lo relativo a la edad promedio de los datos.

En el apéndice 2-A de este capítulo se deducen la media y la variación del error de pronóstico para los promedios móviles y el suavizamiento exponencial en términos de la variación de cada observación individual, suponiendo que el proceso de demanda subyacente es estacionario. Mostramos que ambos métodos son imparciales; esto es, el valor esperado del error de pronóstico es cero. Es más, al equiparar las expresiones para las variaciones del error de pronóstico, se obtiene la misma relación entre α y N como elevar al cuadrado la edad promedio de datos. Esto significa que, si se usan el suavizamiento exponencial y los promedios móviles para predecir el mismo patrón de demanda estacionaria, los errores de pronóstico se distribuyen normalmente; asimismo $\alpha = 2/(N + 1)$, por lo que ambos métodos tendrán exactamente la misma distribución de errores de pronóstico. (Sin embargo, esto no significa que los pronósticos obtenidos por los dos métodos sean iguales.)

Pronósticos de varios pasos adelante

Hasta ahora hemos hablado sólo de los pronósticos de un paso adelante. Es decir, hemos asumido que el pronóstico en el periodo t es para la demanda en el periodo $t + 1$. Sin embargo, hay casos donde se busca hacer un pronóstico para más de un paso adelante. Por ejemplo, la planeación de un minorista para la temporada navideña puede requerir de un pronóstico para las ventas decembrinas en junio a fin de tener suficiente tiempo para prepararse. Como el modelo subyacente asumido tanto para los promedios móviles como para el suavizamiento exponencial es estacionario (es decir, que no cambia con el tiempo), los pronósticos de un paso

adelante y de varios pasos adelante para promedios móviles y suavizamiento exponencial son los mismos. Esto es, un pronóstico realizado en junio para las ventas de julio es el mismo que el pronóstico en junio para las ventas de diciembre. (En el caso de un minorista, el supuesto de "estacionario" probablemente esté equivocado, porque es factible que las ventas de diciembre sean mayores que las de un mes normal. Eso sugiere que dichos métodos pueden no ser apropiados en este caso.)

Comparación entre suavizamiento exponencial y promedios móviles

Existen varias similitudes y varias diferencias entre el suavizamiento exponencial y los promedios móviles.

Similitudes

1. Ambos métodos se deducen con el supuesto de que el proceso de demanda subyacente es estacionario (esto es, puede ser representado por una constante más una fluctuación aleatoria con una media igual a cero). Sin embargo, debemos tener en mente que, aunque los métodos son apropiados para series de tiempo estacionarias, no creemos que las series sean estacionarias para siempre. Al ajustar los valores de N y α podemos hacer que los dos métodos sean más o menos sensibles a los cambios en el patrón subyacente de los datos.

2. Ambos métodos dependen de la especificación de un solo parámetro. Para los promedios móviles el parámetro es N, el número de periodos en el promedio móvil, y para el suavizamiento exponencial el parámetro es α, la constante de suavizamiento. Los valores pequeños de N o los valores grandes de α resultan de pronósticos que asignan mayor peso a los datos actuales; valores grandes de N y valores pequeños de α resultan de asignar pesos mayores a los datos pasados. En caso de N pequeños y α grandes puede presentarse una mayor sensibilidad a los cambios en el proceso de demanda, pero producirá errores de pronóstico con mayor varianza.

3. Ambos métodos se retrasarán con respecto a una tendencia, si es que ésta existe.

4. Cuando $\alpha = 2/(N + 1)$, ambos métodos tienen la misma distribución de pronóstico de error. Esto significa que deben tener más o menos el mismo nivel de exactitud, aunque no necesariamente los mismos pronósticos.

Diferencias

1. El pronóstico de suavizamiento exponencial es un promedio ponderado de *todos* los datos pasados (siempre que la constante de suavizamiento sea estrictamente menor que 1). El pronóstico con promedios móviles es un promedio ponderado de sólo los últimos periodos N de datos. Esto puede ser una ventaja importante para los promedios móviles. Un valor atípico (una observación que no es representativa de la población de la muestra) se elimina del pronóstico por promedios móviles después de los N periodos, pero permanece para siempre en el pronóstico de suavizamiento exponencial.

2. Para poder utilizar los promedios móviles, debemos guardar todos los N datos pasados. Con el fin de utilizar el suavizamiento exponencial, sólo necesitamos guardar el último pronóstico. Esta es la ventaja más significativa del método de suavizamiento exponencial, y explica su popularidad en la práctica. Para apreciar lo que resulta de esta diferencia, consideremos un sistema en el que la demanda de 300 000 artículos de inventario se pronostica cada mes utilizando un promedio móvil de 12 meses. Sólo el módulo de pronóstico requiere que se conserven $300\,000 \times 12 = 3\,600\,000$ piezas de información. Si se utilizara el suavizamiento exponencial, sólo sería necesario conservar 300 000 piezas de información. Esta cuestión es menos importante hoy de lo que era antes, de la misma forma que el costo de almacenaje de información ha disminuido enormemente en los años recientes. De cualquier manera, sigue siendo más sencillo manejar un sistema que requiere menos datos. Es principalmente por esta razón que el suavizamiento exponencial parece ser más popular que los promedios móviles para las aplicaciones de planeación de producción.

Problemas para la sección 2.7

22. Handy, Inc. produce una calculadora electrónica de energía solar que ha experimentado el siguiente historial mensual de ventas durante los primeros cuatro meses del año, en miles de unidades:

Enero	23.3	Marzo	30.3
Febrero	72.3	Abril	15.5

a) Si el pronóstico para enero era 25, determine los pronósticos de un paso adelante de febrero a mayo usando el suavizamiento exponencial con una constante de suavizamiento de $\alpha = .15$.

b) Repita el cálculo del inciso *a* para obtener un valor de $\alpha = .40$. ¿Qué diferencia observa en los pronósticos?

c) Calcule los ECM para los pronósticos que obtuvo en los incisos *a* y *b* desde febrero hasta abril. ¿Qué valor de α dio pronósticos más exactos, basados en el ECM?

23. Compare y contraste el suavizamiento exponencial cuando α es pequeña (cerca de cero) y cuando α es grande (cerca de 1).

24. Las ventas semanales de martillos observadas en la ferretería local en un periodo de ocho semanas fueron de 14, 9, 30, 22, 34, 12, 19, 23.

a) Suponga que se utilizan promedios móviles de tres semanas para pronosticar ventas. Determine el pronóstico de un solo paso para las semanas 4 a 8.

b) Suponga que se utiliza suavizamiento exponencial con una constante de suavizamiento de $\alpha = .15$. Encuentre los pronósticos de suavizamiento exponencial para las semanas 4 a 8. [Para iniciar el método, utilice el mismo pronóstico para la semana 4 que utilizó en el inciso *a*).]

c) Con base en la DAM, ¿qué método funcionó mejor?

d) ¿Cuál es el pronóstico de suavizamiento exponencial al final de la semana 6, para las ventas de la semana 12?

25. Determine lo siguiente:

a) El valor de α consistente con $N = 6$ en promedios móviles.

b) El valor de N consistente con $\alpha = .05$.

c) El valor de α que resulta en una varianza de pronóstico de error σ_e^2, 10% mayor a la variación de cada observación, σ^2 (consulte las fórmulas obtenidas en el apéndice 2-A).

26. En relación con los datos del problema 22, ¿cuál es el pronóstico de suavizamiento exponencial realizado al final de marzo para las ventas en julio? Suponga que $\alpha = .15$.

27. Para los datos de los problemas 16 a 21, use el promedio aritmético de los primeros seis meses de datos como base para inicializar el suavizamiento exponencial.

a) Determine los pronósticos de suavizamiento exponencial un paso adelante desde agosto hasta diciembre, suponiendo que $\alpha = .20$.

b) Compare la exactitud de los pronósticos obtenidos en el inciso *a*) con los determinados en el problema 20, con promedio móvil de seis meses y de un paso adelante.

c) Comente las razones por las que obtuvo el resultado del inciso *b*).

Aplicación instantánea

SPORT OBERMEYER REDUCE SUS COSTOS CON MEJORES PRONÓSTICOS[1]

Sport Obermeyer es un proveedor líder en el mercado estadounidense de ropa para esquiar. La compañía, fundada en 1950 por el instructor de esquí e ingeniero Klaus Obermeyer, disfruta de cerca de 45% de la participación del mercado en el rápidamente creciente mercado infantil y 11% de la participación del mercado adulto. Virtualmente todas las ofertas de la compañía se rediseñan anualmente para incorporar cambios en estilo, telas y colores. Durante más de 30 años, la compañía pudo cubrir con éxito las demandas al dedicarse a la producción durante los meses del verano después de recibir los pedidos de los clientes.

Durante la década pasada, las cosas cambiaron y surgieron problemas. En primer lugar, se incrementó el volumen. Hubo capacidad insuficiente entre los proveedores para generar el volumen requerido en el verano. En segundo lugar, la compañía desarrolló una compleja estrategia de cadena de abastecimiento global (véase la sección 10 del capítulo 6) para reducir costos. Una *parka* vendida en Estados Unidos podía coserse en China con tela y partes de Japón, Corea del Sur y Alemania. Todos estos cambios juntos alargaron el tiempo de demora en producción, por lo que la compañía tuvo que comprometerse a producir antes de que llegaran los pedidos por parte de los clientes.

La compañía emprendió varias iniciativas de “respuesta rápida” para reducir los tiempos libres. Éstas incluyeron animar a algunos clientes a hacer sus pedidos con mayor anticipación, localizar materias primas cerca de sus instalaciones del Lejano Oriente, e instituir un sistema de flete aéreo para acelerar la entrega del Lejano Oriente a su centro de distribución en Detroit. A pesar de estos cambios, no se pudo resolver el problema de falta de inventario y descuentos debido al exceso de abasto. La compañía tenía que comprometer todavía la mitad de la producción con base en los pronósticos. En la industria del vestido, es común que no exista un historial estadístico sobre el cual se puedan hacer pronósticos, y los errores de pronóstico pueden ser muchos. Son muy comunes los productos que se venden más que lo originalmente pronosticado por un factor de 2 o que quedan por debajo de los pronósticos originales por un factor de 10.

Sport Obermeyer necesitó apoyarse en el pronóstico para evitar errores de cálculo costosos. El procedimiento acostumbrado era basar los pronósticos en un consenso de miembros del comité de compras. El problema con el pronóstico de consenso es que las personalidades dominantes en un grupo tienen un mayor peso. Un pronóstico obtenido en esta forma puede representar sólo la opinión de una persona. Para superar este problema, el equipo de investigación (Fisher *et. al*., 1994) recomendó que los miembros del comité emitieran pronósticos individuales.

La dispersión entre los pronósticos individuales pareció ser un indicador confiable de la exactitud del pronóstico. Cuando los pronósticos de los distintos miembros del comité eran parecidos, solían ser más exactos. Esto proporcionó un mecanismo para señalar los productos cuyas ventas parecían tener un pronóstico deficiente. El problema de un pronóstico deficiente de los artículos no se resolvió con esto, pero permitió a la compañía comprometerse primero con la producción de artículos cuyos pronósticos tenían mayores posibilidades de ser exactos. Para el momento en que la producción tenía que empezar con los artículos problemáticos podría contarse con patrones tempranos de ventas.

El equipo se dio cuenta de que los minoristas compartían muchas similitudes. Eso significó que incluso si sólo se recibía el primer 20% de los pedidos para un producto, la información podría mejorar espectacularmente los pronósticos. Los planes de producción para estos artículos “problemáticos” estarían ahora comprometidos con mayor confianza. De esta manera, la compañía pudo separar los productos en dos categorías: reactivos y no reactivos. Los artículos no reactivos son aquellos para los cuales el pronóstico puede ser exacto. Éstos se producen al principio de la temporada a partir de las cifras de ventas tempranas. La experiencia de la firma fue que los costos de desabasto y descuento se redujeron de 10.2% de las ventas a 1.8% de las ventas en artículos que podrían producirse reactivamente. Sport Obermeyer fue capaz de producir el 30% de su volumen de ventas de manera reactiva y experimentó una reducción de costo de aproximadamente 5% de las ventas.

¿Cuáles son las lecciones aquí? Una es que, incluso en los casos en los que no existe un historial estadístico, la metodología estadística puede aplicarse con mucho éxito para mejorar la exactitud del pronóstico. Otra es no suponer que las cosas deben hacerse de cierta manera. Sport Obermeyer asumió que el pronóstico de consenso fue el mejor enfoque. De hecho, al requerir que el comité de compras llegara a un consenso, se estaba ignorando información muy valiosa. Las diferencias entre pronósticos individuales probaron ser muy importantes.

---

[1] Esta aplicación se basa en el trabajo de un equipo de la Wharton School y de la Harvard Business School. Los resultados se reportaron en Fisher *et al*. (1994).

2.8 MÉTODOS BASADOS EN LA TENDENCIA

Los pronósticos de suavizamiento exponencial y de promedios móviles se retrasarán en relación con una tendencia en caso de que la haya. Consideraremos dos métodos de pronóstico que representan específicamente una tendencia en los datos: el análisis de regresión y el método de Holt. El análisis de regresión es un método que ajusta una línea recta a un conjunto de datos. El método de Holt es un tipo de suavizamiento exponencial doble que permite un suavizamiento simultáneo en la serie y en la tendencia.

Análisis de regresión

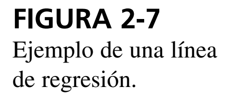
Sean (x_i, y_1), (x_2, y_2), . . . , (x_n, y_n) n datos apareados para las dos variables X y Y. Supongamos que y_i es el valor observado de Y cuando x_i es el valor observado de X. Refiérase a Y como la variable dependiente y a X como la variable independiente. Creemos que existe una relación entre X y Y que puede representarse mediante una línea recta.

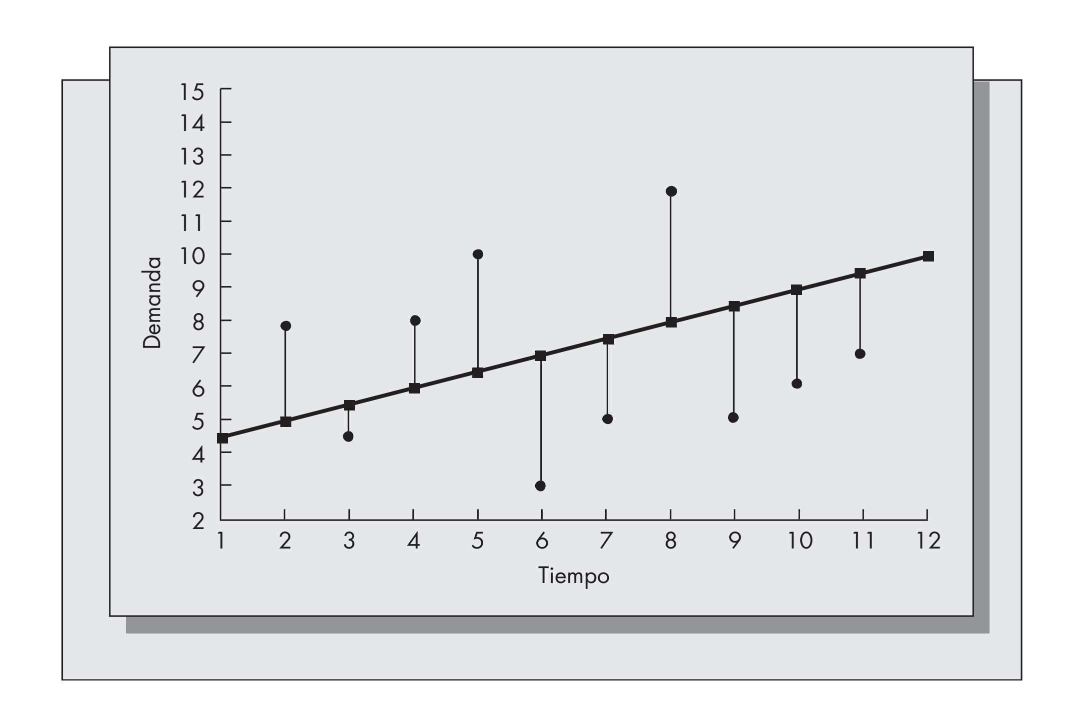
$$\hat{Y} = a + bX.$$

Interpretemos $\hat{Y}$ como el valor predicho de Y. La meta es encontrar los valores de a y b de manera que la recta $\hat{Y} = a + bX$ proporcione el mejor ajuste de los datos. Los valores de a y b se eligen de manera que se minimice la suma de las distancias cuadráticas entre la línea de regresión y los puntos de datos (véase la figura 2-7). En el apéndice 2-B se deducen los valores óptimos de a y b en términos de los datos dados.

Cuando aplicamos el análisis de regresión al problema de pronóstico, la variable dependiente muchas veces corresponde al tiempo y la variable dependiente a la serie que se va a pronosticar. Supongamos que D_1, D_2, . . . , D_n son los valores de la demanda en los tiempos 1, 2, . . . , n. En el apéndice 2-B se muestra que los valores óptimos de a y b están dados por

$$b = \frac{S_{xy}}{S_{xx}}$$


FIGURA 2-7
Ejemplo de una línea de regresión.



y

$$a = \overline{D} - b(n + 1)/2,$$

donde

$$S_{xy} = n \sum_{i=1}^{n} iD_i - \frac{n(n + 1)}{2} \sum_{i=1}^{n} D_i,$$

$$S_{xx} = \frac{n^2(n + 1)(2n + 1)}{6} - \frac{n^2(n + 1)^2}{4},$$

y $\overline{D}$ es el promedio aritmético de las demandas observadas durante los periodos 1, 2, . . . , n.

Ejemplo *2.4* Aplicaremos el análisis de regresión al problema tratado en los ejemplos 2.2. y 2.3, referentes a predecir fallas en los motores de las aeronaves. Recordemos que la demanda de motores para aeronaves durante los últimos ocho trimestres fue de 200, 250, 175, 186, 225, 285, 305, 190. Supongamos que usamos los primeros cinco periodos como base para poder estimar los parámetros de regresión. Entonces,

$$\begin{aligned} S_{xy} &= 5[200 + (250)(2) + (175)(3) + (186)(4) + (225)(5)] \\ &\quad - [(5)(6)/2][200 + 250 + 175 + 186 + 225] \\ &= -70, \\ S_{xx} &= (25)(6)(11)/6 - (25)(36)/4 = 50. \end{aligned}$$

Así,

$$\begin{aligned} b &= S_{xy}/S_{xx} = -70/50 = -7/5, \\ a &= 207.2 - (-27/5)(3) = 211.4. \end{aligned}$$

Se deduce que la ecuación de regresión basada en cinco periodos de datos es

$$\hat{D}_t = 211.4 - (7/5)t.$$

$\hat{D}_t$ es el valor predicho de la demanda en el tiempo t. Emplearemos esta ecuación de regresión para pronosticar del periodo 5 a cualquier periodo posterior a éste. Por ejemplo, el pronóstico hecho en el periodo 5 para el periodo 8 se obtendrá al sustituir $t = 8$ en la ecuación de regresión que se acaba de dar, lo que resultará en el pronóstico 211.4 − (7/5)(8) = 200.2. Observemos que si estuviéramos interesados en hacer pronósticos en el periodo 7 para el periodo 8, entonces esta ecuación de regresión no sería adecuada. Debemos repetir todo el cálculo usando los datos de los periodos 1 a 7. De hecho, uno de los inconvenientes más serios de utilizar la regresión para el pronóstico es que resulta muy engorroso actualizar los pronósticos conforme se va contando con nuevos datos.

Problemas para la sección 2.8

28. Shoreline Park en Mountain View, California, ha mantenido registros del número de clientes habituales que visitan el parque desde su apertura en enero de 1993. Durante los primeros seis meses de operación, se registraron las siguientes cifras:

Mes	Número de clientes	Mes	Número de clientes
Enero	133	Abril	640
Febrero	183	Mayo	1 876
Marzo	285	Junio	2 550

a) Dibuje una gráfica de estos seis datos. Suponga que enero = periodo 1, febrero = periodo 2, etc. Con una regla, trace "a simple vista" la recta que mejor se ajuste a los datos. Calcule la pendiente y la intercepción en su gráfica.

b) Calcule los valores exactos de la intercepción a y la pendiente b a partir de las ecuaciones de regresión.

c) ¿Cuáles son los pronósticos obtenidos de julio a diciembre de 1993 con la ecuación de regresión determinada en el inciso *b*)?

d) Comente los resultados que obtuvo en el inciso *c*). Específicamente, ¿cuánta confianza debe tenerse en la exactitud de los pronósticos que obtuvo?

29. El departamento de parques y actividades recreativas de Mountain View debe proyectar el uso total de Shoreline Park para el año 1994.

a) Determine el pronóstico para el número total de personas que usarán el parque en 1994 de acuerdo con la ecuación de regresión.

b) Determine el pronóstico para la misma cantidad usando un promedio móvil de seis meses.

c) Dibuje la gráfica más atinada de la curva que describa el uso del parque mes a mes durante el año civil, y pronostique la misma cantidad usando su gráfica. ¿Se acerca su predicción a la respuesta que obtuvo en el inciso *a*) o el inciso *b*)? Describa sus resultados.

Suavizamiento exponencial doble usando el método de Holt

El método de Holt es un tipo de suavizamiento exponencial doble diseñado para dar seguimiento a series de tiempo con tendencia lineal. El método requiere de la especificación de dos constantes de suavizamiento, α y β, y utiliza dos ecuaciones de suavizamiento: una para el valor de la serie (la intercepción) y una para la tendencia (la pendiente). Las ecuaciones son

$$S_t = \alpha D_t + (1 - \alpha)(S_{t-1} + G_{t-1}),$$
$$G_t = \beta(S_t - S_{t-1}) + (1 - \beta)G_{t-1}.$$

Interprete S_t como el valor de la intercepción en el tiempo t y G_t como el valor de la pendiente en el tiempo t. La primera ecuación es muy similar a la utilizada para el suavizamiento exponencial simple. Cuando la observación de la demanda más común, D_t, está disponible, se promedia con el pronóstico previo de la demanda actual, que es la intercepción previa, S_{t-1}, más 1 vez la pendiente previa, G_{t-1}. La segunda ecuación puede explicarse de la siguiente manera: nuestro nuevo estimado de intercepción, S_t, provoca que modifiquemos nuestro estimado de pendiente en la cantidad $S_t - S_{t-1}$. Este valor se promedia entonces con el estimado anterior de la pendiente G_{t-1}. Las constantes de suavizamiento pueden ser las mismas, pero para la mayoría de las aplicaciones se da mayor estabilidad al estimado de la pendiente (lo que implica que $\beta \leq \alpha$).

El pronóstico de τ pasos adelante, hecho en el periodo t, que se denota como $F_{t,\, t+\tau}$, está dado por

$$F_{t,\, t+\tau} = S_t + \tau G_t.$$

Ejemplo *2.5*

Apliquemos el método de Holt al problema de desarrollar datos de pronósticos de un paso adelante para las fallas de motores de aeronaves. Recuerde que la serie original fue 200, 250, 175, 186, 225, 285, 305, 190. Supongamos que tanto α como β son iguales a 0.1. Para iniciar el método, necesitamos estimados tanto de la intercepción como de la pendiente en el tiempo cero. Supongamos que éstas son $S_0 = 200$ y $G_0 = 10$. Entonces obtenemos

$$S_1 = (.1)(200) + (.9)(200 + 10) = 209.0$$
$$G_1 = (.1)(209 - 200) + (.9)(10) = 9.9$$

$$S_2 = (.1)(250) + (.9)(209 + 9.9) = 222.0$$

$$G_2 = (.1)(222 - 209) + (.9)(9.9) = 10.2$$

$$S_3 = (.1)(175) + (.9)(222 + 10.2) = 226.5$$

$$G_3 = (.1)(226.5 - 222) + (.9)(10.2) = 9.6$$

y así sucesivamente.

Al comparar los pronósticos de un paso adelante con el número real de fallas para los periodos 4 a 8, obtenemos lo siguiente:

Periodo	Real	Pronóstico	|Error|
4	186	236.1	50.1
5	225	240.3	15.3
6	285	247.7	37.3
7	305	260.8	44.2
8	190	275.0	85.0

Al promediar los números de la última columna, obtenemos una DAM de 46.4. Observe que éste es menor que en los casos de promedios móviles o suavizamiento exponencial. El método de Holt funciona mejor para esta serie porque está particularmente diseñado para dar seguimiento a la tendencia en los datos, no así el suavizamiento exponencial simple y los promedios móviles. Observe que los pronósticos que aparecen en la tabla son pronósticos de un paso adelante. Supongamos que usted deseaba pronosticar la demanda en el periodo 2 para el periodo 5. Este pronóstico es $F_{2,\,5} = S_2 + (3)G_2 = 222 + (3)(10.2) = 252.6$.

El problema de inicialización surge también al comenzar con el método de Holt. El mejor enfoque es establecer un cierto conjunto de periodos iniciales como base y utilizar el análisis de regresión para determinar los estimados de la pendiente y la intercepción usando los datos de base.

Tanto el método de Holt como la regresión están diseñados para manejar series que muestran tendencia. Sin embargo, con el método de Holt no resulta sencillo en absoluto actualizar pronósticos conforme se obtienen nuevas observaciones.

Más problemas para la sección 2.8

30. Para los datos del problema 28, utilice los resultados de la ecuación de regresión para determinar la pendiente e intercepción de la serie a finales de junio. Use esos números como los valores iniciales de pendiente e intercepción que requiere el método de Holt. Suponga que $\alpha = .15$, $\beta = .10$ para todos los cálculos.

    a) Suponga que el número real de visitantes que acudieron al parque durante julio fue de 2 150 y durante agosto fue de 2 660. Utilice el método de Holt para actualizar los estimados de pendiente e intercepción con base en estas observaciones.

    b) ¿Cuáles son los pronósticos de un paso adelante y dos pasos adelante que da el método de Holt para el número de visitantes al parque en septiembre y octubre?

    c) ¿Cuál es el pronóstico hecho al final de julio para el número de asistentes al parque durante diciembre?

31. Debido a serias inundaciones, el parque se cerró la mayor parte de diciembre de 1993. Durante ese tiempo únicamente recibió la visita de 53 personas. Comente los efectos que pudo haber tenido esta observación en las predicciones del uso futuro del parque. Si usted estuviera a cargo de realizar los pronósticos de uso del parque, ¿cómo manejaría este dato?

32. Describa algunos de los problemas que pudieran surgir al usar el análisis de regresión o el método de Holt para obtener pronósticos de varios pasos adelante.

2.9 MÉTODOS PARA SERIES ESTACIONALES

El último tema de este capítulo se refiere a los métodos de pronóstico para problemas estacionales. Una serie estacional es aquella que tiene un patrón que se repite cada *N* periodos, para algún valor de *N* (que es cuando menos 3). Una serie estacional típica se muestra en la figura 2-8.

Llamaremos "duración de la estación" al número de periodos antes de que el patrón comience a repetirse (*N* en la figura). Observe que es diferente del uso popular de la palabra *estación* como una temporada del año. Para poder emplear un modelo estacional, debemos ser capaces de especificar cuánto dura una estación.

Hay varias formas de representar la estacionalidad. La más común es suponer que existe un conjunto de multiplicadores c_t, para $1 \leq t \leq N$, con la propiedad de que $\sum c_t = N$. El multiplicador c_t representa la cantidad promedio que la demanda en el periodo *t* de la estación está por encima o por debajo del promedio global. Por ejemplo, si $c_3 = 1.25$ y $c_5 = .60$, entonces, en promedio, la demanda en el tercer periodo de la estación está 25% por encima de la demanda promedio y la demanda en el quinto periodo de la estación es 40% menor a la demanda promedio. Estos multiplicadores se conocen como factores estacionales.

Factores estacionales para series estacionarias

En esta parte de la sección presentamos un método sencillo para calcular los factores estacionales para una serie de tiempo con variación estacional y sin tendencia. En la siguiente parte de esta sección consideramos un método que parece ser más exacto cuando también se presenta una tendencia. Ambos métodos requieren un mínimo de dos estaciones de datos.

El método es el siguiente:

1. Calcule la media de la muestra de todos los datos.
2. Divida cada observación por la media de la muestra. Esto da los factores estacionales para cada periodo de datos observados.
3. Promedie los factores para los periodos semejantes dentro de cada estación. Esto es, promedie todos los factores correspondientes al primer periodo de una estación, todos los facto-

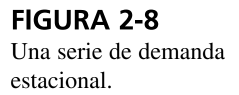
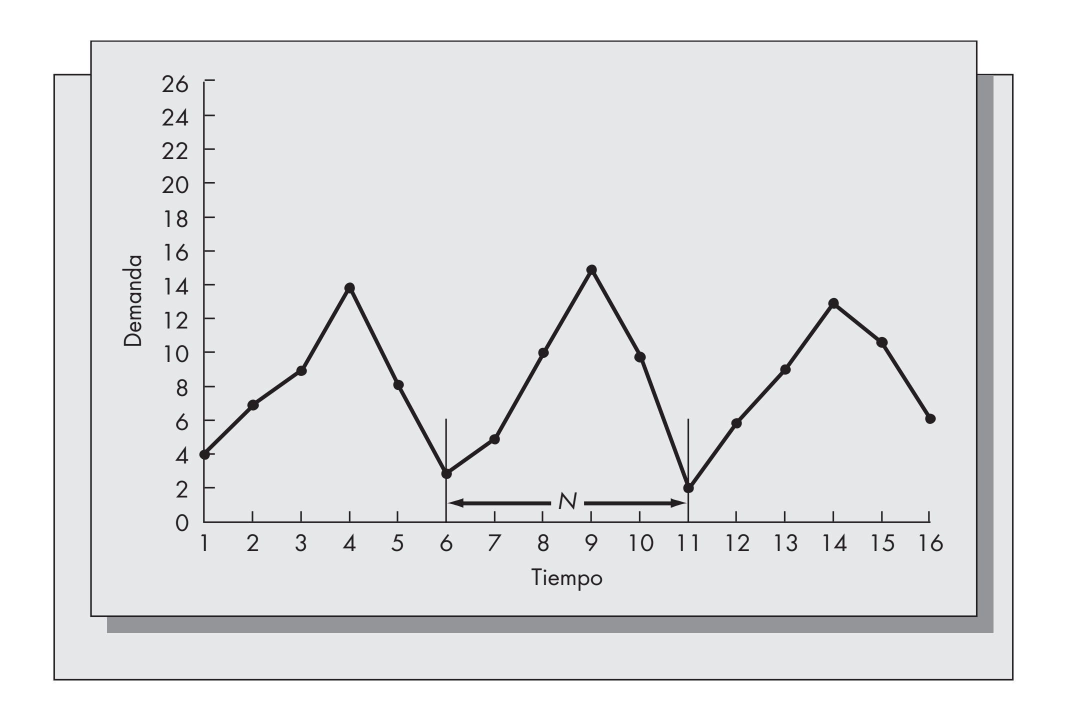


FIGURA 2-8
Una serie de demanda estacional.



res correspondientes al segundo periodo de una estación, y así sucesivamente. Los promedios resultantes son los N factores estacionales. Siempre sumarán exactamente N.

Ejemplo ***2.6*** El departamento de transporte de una ciudad necesita determinar los factores de uso, por día de la semana, para un popular puente de cuota que conecta diferentes partes de la ciudad. En el estudio actual sólo se contemplan los días hábiles. Supongamos que el número de coches que utilizaron el puente cada día hábil durante las pasadas cuatro semanas fue (en miles de automóviles)

	Semana 1	**Semana 2**	**Semana 3**	**Semana 4**
Lunes	16.2	17.3	14.6	16.1
Martes	12.2	11.5	13.1	11.8
Miércoles	14.2	15.0	13.0	12.9
Jueves	17.3	17.6	16.9	16.6
Viernes	22.5	23.5	21.9	24.3

Encuentre los factores estacionales correspondientes al uso diario del puente.

Solución

1. Primero calculamos el promedio de todas las observaciones. El resultado debe ser 16.425.
2. Después dividimos cada observación entre la media de la muestra. Al hacerlo se obtiene lo siguiente:

0.986	1.053	0.889	0.980
0.743	0.700	0.798	0.718
0.865	0.913	0.791	0.785
1.053	1.072	1.029	1.011
1.370	1.431	1.333	1.479

3. Finalmente, promediamos los factores correspondientes al mismo periodo de la estación. Esto es, promediamos todos los factores para los lunes, todos los factores para los martes, y así sucesivamente. Los cinco factores estacionales resultantes para uso del puente por día de la semana son

Lunes:	0.98
Martes:	0.74
Miércoles:	0.84
Jueves:	1.04
Viernes:	1.40

La suma de estos factores debe ser 5, como se requirió. Interprete los resultados en la siguiente forma: los lunes son días promedio porque el factor para éstos está muy cerca de 1. El uso es menor los martes, cerca de 26% por debajo del promedio, los viernes son los días más altos, cerca de 40% por encima del promedio.

Los pronósticos para el número de autos que utilizan el puente en un día de la semana se obtendrá mediante la multiplicación de la media de la muestra que es 16.425 por el factor estacional apropiado. Por lo tanto, los pronósticos de uso por día de la semana (en miles de autos) será

Lunes:	16.1
Martes:	12.1
Miércoles:	13.8
Jueves:	17.1
Viernes:	23.0

Descomposición estacional usando promedios móviles

Un método ligeramente más complejo para estimar los factores estacionales requiere del cálculo de los promedios móviles de *N* periodos, donde *N* es la duración de la estación. El método se ilustra mejor a través de un ejemplo.

Ejemplo *2.7*

Supongamos que el historial de demanda original de cierto artículo en los últimos ocho trimestres es 10, 20, 26, 17, 12, 23, 30, 22. La gráfica de esta serie de demanda se presenta en la figura 2-9.

La imagen sugiere que estos datos representan dos estaciones, en las que cada una tiene cuatro periodos. El siguiente paso es calcular los promedios móviles de esos cuatro periodos. Esto da como resultado

Periodo	Demanda	PM(4)
1	10	
2	20	
3	26	
4	17	18.25
5	12	18.75
6	23	19.50
7	30	20.50
8	22	21.75

Después, los promedios móviles deben "centrarse". El promedio móvil 18.25 se calcula al promediar la demanda de los periodos 1, 2, 3 y 4. El centro de estos números es 2.5. El siguiente promedio de 18.75 corresponde a los periodos 2, 3, 4 y 5. El centro de estos números es 3.5. Repetimos este

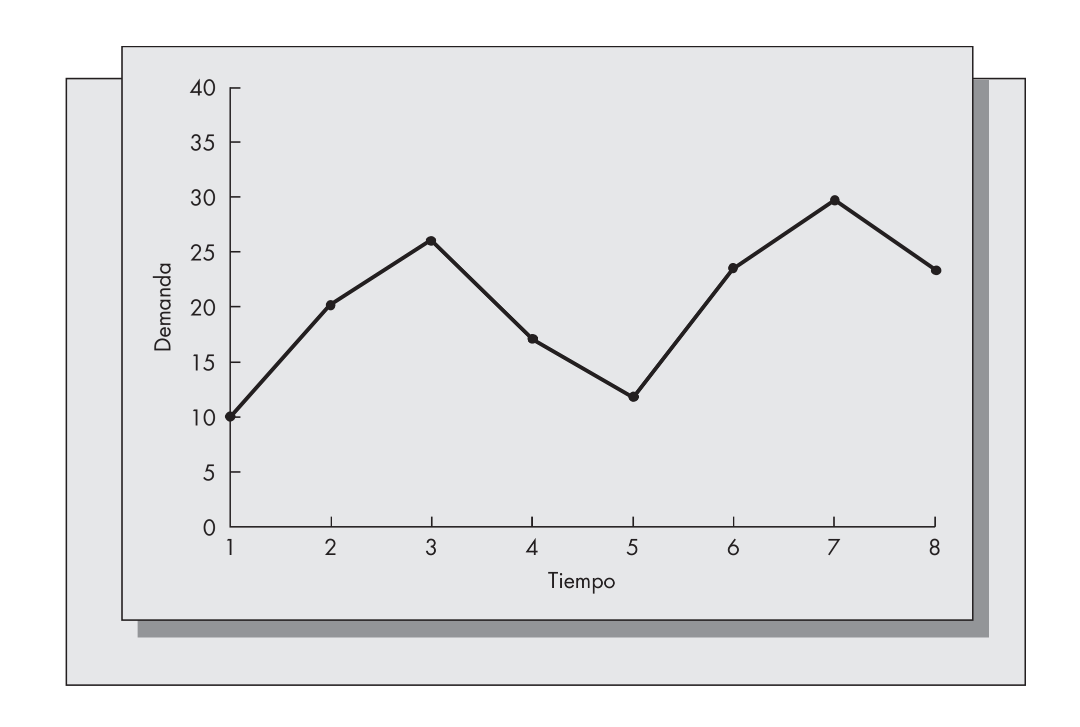


FIGURA 2-9 Historia de la demanda para el ejemplo 2.7.

proceso con los promedios restantes y los ubicamos en las posiciones centradas de la siguiente manera:

Periodo	Demanda	PM(4)	Centrado
1	10		
2	20		
			18.25
3	26		
			18.75
4	17	18.25	
			19.50
5	12	18.75	
			20.50
6	23	19.50	
			21.75
7	30	20.50	
8	22	21.75	

Si N hubiera sido impar en lugar de par, los valores centrados coincidirían con los periodos en vez de estar entre periodos. [Por ejemplo, para $N = 5$ el primer valor PM(5) está centrado en el periodo 3].

Enseguida, debemos hacer que estos valores vuelvan a coincidir con los periodos. Para hacer esto, promediamos los valores adyacentes. Al promediar 18.25 y 18.75 se obtiene 18.5, que corresponde al promedio de periodos 2.5 y 3.5, que es 3. Al repetir este proceso para los otros valores centrados también se obtienen promedios móviles centrados para los periodos 4, 5 y 6.

El siguiente paso consiste en hallar los valores para los periodos 1, 2, 7 y 8. Para obtener los valores de los periodos 1 y 2, promediamos los valores para los periodos 3 y 4 y, para obtener los valores de los periodos 7 y 8, promediamos los valores 5 y 6. De esta forma, obtenemos un promedio móvil centrado para cada periodo en el que tenemos una observación de demanda.

(A) Periodo	(B) Demanda	(C) PM centrado	(B/C) Razón
1	10	18.81	0.532
2	20	18.81	1.063
3	26	18.50	1.405
4	17	19.125	0.888
5	12	20.00	0.600
6	23	21.125	1.089
7	30	20.56	1.463
8	22	20.56	1.070

Una vez que hemos obtenido el promedio móvil centrado para cada periodo, formamos la razón de la demanda para ese periodo sobre el PM centrado. Estos valores, reportados en la cuarta columna de la tabla anterior, son estimados de los factores estacionales para cada periodo. El siguiente paso es formar el promedio de los factores correspondientes a los mismos periodos de cada estación. Como tenemos dos estaciones de datos, promediamos $(0.532 + 0.600)/2 = 0.566$, $(1.063 + 1.089)/2 = 1.076$, $(1.405 + 1.463)/2 = 1.434$, y $(0.888 + 1.070)/2 = 0.979$. En este caso, el procedimiento rinde exactamente cuatro factores estacionales. En general, da como resultado exactamente N factores estacionales.

Finalmente, debemos estar seguros de que la suma de los factores estacionales sea exactamente N, o exactamente 4 en el caso de este ejemplo. Encontramos 0.566 + 1.076 + 1.434 + 0.979 = 4.055. Para poder hacer la suma exactamente igual a 4, multiplicamos cada factor por 4/4.055 = 0.9864. Los factores estacionales finales son

Periodo	Factor
1	0.558
2	1.061
3	1.415
4	0.966

La suma de estos factores estacionales finales es exactamente 4. Con base en los datos proporcionados, estos factores indican que, en promedio, el primer trimestre de cada año da como resultado ventas que son aproximadamente 45% menores al promedio anual, el segundo trimestre en ventas está más o menos 6% por encima del promedio anual, etcétera. El siguiente paso en el proceso es dividir cada observación por el factor estacional apropiado para poder obtener la demanda desestacionalizada.

(A) Demanda	(B) Factor	Demanda desestacionalizada (A/B)
10	0.558	17.92
20	1.061	18.85
26	1.415	18.39
17	0.966	17.60
12	0.558	21.50
23	1.061	21.68
30	1.415	21.22
22	0.966	22.77

La serie desestacionalizada seguirá conteniendo todos los componentes de la señal de la serie original, excepto la estacionalidad. Puede hacerse ahora un pronóstico con base en la demanda desestacionalizada, que debe "reestacionalizarse" mediante la multiplicación del factor estacional apropiado.

Como ejemplo, supongamos que estamos usando un promedio móvil de seis meses para pronosticar la serie desestacionalizada. El promedio de las últimas seis observaciones es 20.52. Este número se multiplica entonces por el factor estacional apropiado para obtener los pronósticos de demanda futura. El pronóstico para el periodo 9 es (20.52)(0.5580) = 11.45, para el periodo 10 es (20.52)(1.061) = 21.77, y así sucesivamente.

Sin embargo, como la serie desestacionalizada muestra una tendencia, sería más adecuado aplicarle un método basado en tendencia, como el método de Holt o el análisis de regresión. Si usamos el análisis de regresión, obtenemos el ajuste por mínimos cuadrados de los datos como $\hat{D}_t = 16.8 + .7092t$. Supongamos que nos interesa pronosticar al final del periodo 8 los periodos desde 9 hasta 12. Al sustituir de 9 a 12 en lugar de t en la ecuación de regresión, se obtienen los pronósticos de las series desestacionalizadas para el siguiente año; éstos son 23.18, 23.89, 24.60 y 25.31. Cada uno de estos pronósticos se vuelve a estacionalizar multiplicándolo por el factor estacional apropiado. Los pronósticos obtenidos por este método para el año siguiente (periodos 9 a 12) son respectivamente 12.93, 25.35, 34.81 y 24.45.

Problemas para la sección 2.9

33. Las ventas de pantalones cortos para caminata de Hugo's Department Store en el centro de Rolla parece mostrar un patrón estacional. El propietario de la tienda, Wally Hugo, ha mantenido cuidadosos registros de las ventas de varios de sus populares artículos, incluyendo los pantalones cortos para caminata.

Durante los últimos dos años, las ventas mensuales de los pantalones cortos han sido

	Año 1	Año 2		Año 1	Año 2
Enero	12	16	Julio	112	130
Febrero	18	14	Agosto	90	83
Marzo	36	46	Septiembre	66	52
Abril	53	48	Octubre	45	49
Mayo	79	88	Noviembre	23	14
Junio	134	160	Diciembre	21	26

Suponiendo que no se ha presentado ninguna tendencia en las ventas de pantalones cortos durante los últimos dos años, obtenga los estimados para los factores estacionales de cada mes.

34. Una popular marca de zapatos deportivos cuenta con el siguiente historial de demanda por trimestres en un periodo de tres años.

1990	Demanda	1991	Demanda	1992	Demanda
1	12	1	16	1	14
2	25	2	32	2	45
3	76	3	71	3	84
4	52	4	62	4	47

a) Determine los factores estacionales de cada trimestre por el método de promedios móviles centrados.

b) Con base en el resultado del inciso *a*, determine la serie de demanda desestacionalizada.

c) Pronostique la demanda en el primer trimestre de 1993 para la serie desestacionalizada de un promedio móvil de seis trimestres.

d) Utilice los resultados de los incisos *a* y *c*, para pronosticar la demanda de calzado en el primer trimestre de 1993.

e) Compare el resultado del inciso *a* con factores estacionales obtenidos por un medio más sencillo descrito al principio de la sección 2.9.

Método de Winters para los problemas estacionales

El método de promedio móvil que se acaba de describir puede utilizarse para predecir una serie estacional con o sin una tendencia. Sin embargo, conforme se van obteniendo nuevos datos, el método requiere que todos los factores estacionales se calculen de nuevo a partir de cero. El método de Winters es un tipo de suavizamiento exponencial triple, y tiene la importante ventaja de ser fácil de actualizar conforme se dispone de nuevos datos.

Suponemos un modelo de la forma

$$D_t = (\mu + G_t)\, c_t + \epsilon_t.$$

Interpretemos μ como la señal base o intercepción en el tiempo $t = 0$ excluyendo la estacionalidad, G_t como el componente de tendencia o pendiente, c_t como el componente estacional multiplicativo en el periodo t, y finalmente ϵ_t como el término de error. Como el factor estacional multiplica tanto al nivel base como al término de tendencia, suponemos que la serie subyacente tiene una forma similar a la que aparece en la figura 2-10.

De nuevo, supongamos que la duración de la estación es exactamente N periodos y los factores estacionales son los mismos cada estación y tienen la propiedad de que $\Sigma\, c_t = N$. Se usan

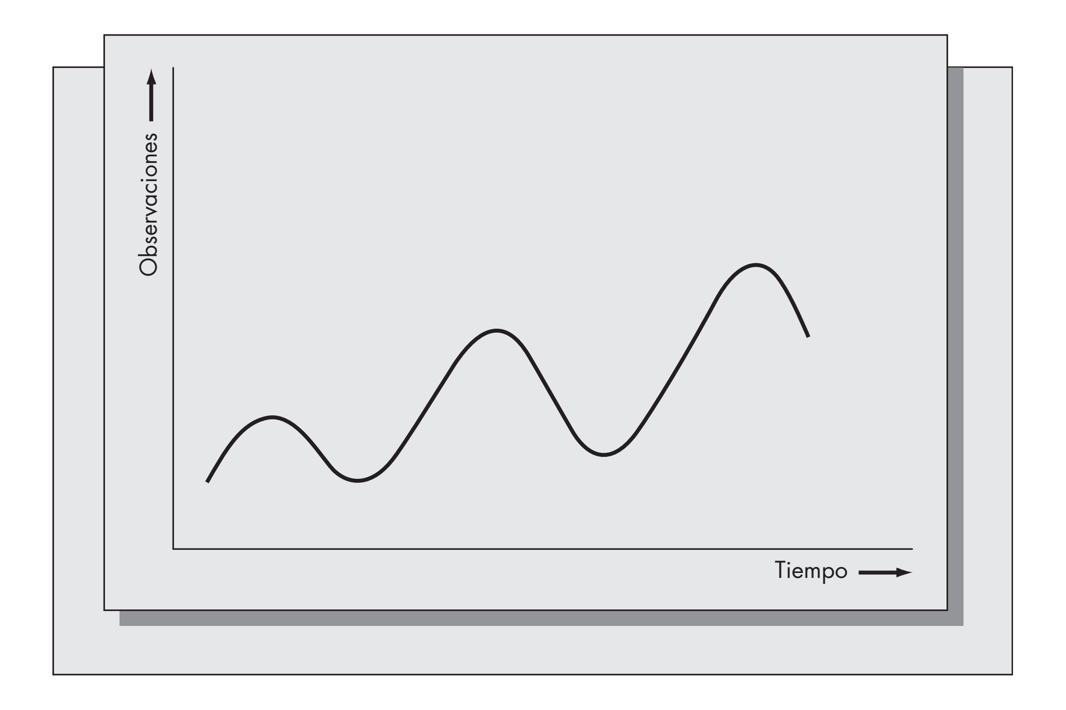


FIGURA 2-10
Serie estacional con tendencia incremental.

tres ecuaciones de suavizamiento en cada periodo para actualizar los cálculos de serie desestacionalizada, los factores estacionales y la tendencia. Estas ecuaciones pueden tener diferentes constantes de suavizamiento, que etiquetaremos como α, β y γ.

1. *La serie*. El nivel actual de la serie desestacionalizada, S_t, está dado por

$$S_t = \alpha(D_t/c_{t-N}) + (1 - \alpha)(S_{t-1} + G_{t-1}).$$

Observe lo que hace esta ecuación. Al dividir el factor estacional apropiado, estamos desestacionalizando la observación de demanda más nueva. Esto se promedia entonces con el pronóstico actual para la serie desestacionalizada, como en el método de Holt.

2. *La tendencia*. La tendencia se actualiza en una forma similar a la del método de Holt.

$$G_t = \beta[S_t - S_{t-1}] + (1 - \beta)G_{t-1}.$$

3. *Los factores estacionales*.

$$c_t = \gamma(D_t/S_t) + (1 - \gamma)c_{t-N}.$$

La relación de la observación de demanda más reciente sobre el estimado actual de la demanda desestacionalizada da como resultado el estimado actual del factor estacional. Después, esto se promedia con el mejor estimado previo del factor estacional, c_{t-N}. Cada vez que se actualiza un factor estacional, es necesario normalizar los N factores más recientes para que sumen N.

Por último, el pronóstico realizado en el periodo t para cualquier periodo futuro $t + \tau$ está dado por

$$F_{t,\, t+\tau} = (S_t + \tau G_t)c_{t+\tau-N}.$$

Obsérvese que esta ecuación de pronóstico supone que $t \leq N$. Si $N < \tau \leq 2N$, el factor estacional adecuado será $c_{t+\tau-2N}$; si $2N < \tau \leq 3N$, el factor estacional adecuado será $c_{t+\tau-3N}$, y así sucesivamente.

Procedimiento de inicialización

Para iniciar el método, necesitamos los estimados iniciales de la serie, la pendiente y los factores estacionales. Winters sugiere que, para la inicialización, deben estar disponibles dos estaciones de datos. Supongamos que están disponibles exactamente dos estaciones de datos; esto es $2N$ datos. Supongamos también que el periodo actual es $t = 0$, así que las observaciones pasadas se marcan como $D_{-2N+1}, D_{-2N+2}, \ldots, D_0$.

1. Calcule por separado la media de la muestra de las dos estaciones de datos.

$$V_1 = \frac{1}{N} \sum_{j=-2N+1}^{-N} D_j$$

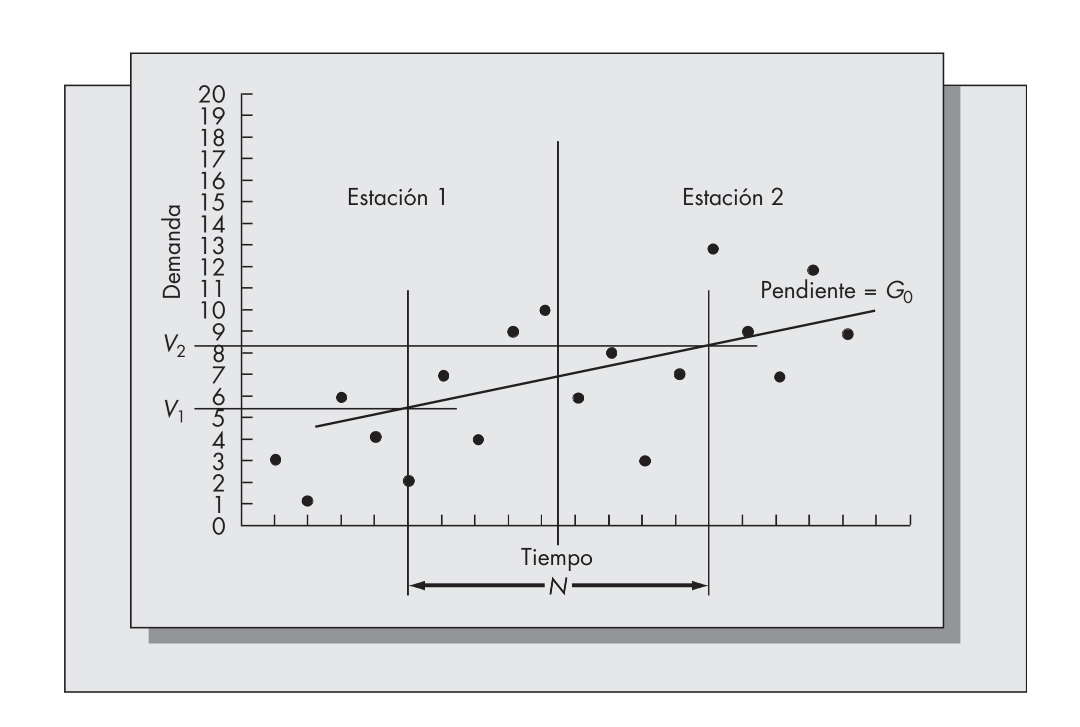
$$V_2 = \frac{1}{N} \sum_{j=-N+1}^{0} D_j$$

2. Defina $G_0 = (V_2 - V_1)/N$ como el estimado de la pendiente inicial. Si hay $m > 2$ estaciones de datos disponibles para la inicialización, entonces calcule $V_1, \ldots, V_m$ como en el paso 1 y defina $G_0 = (V_m - V_1)/[(m - 1)N]$. Si localizamos V_1 en el centro de la primera estación de datos [en el periodo $(-3N + 1)/2$] y V_2 en el centro de la segunda estación de datos [en el periodo $(-N + 1)/2$], entonces G_0 es simplemente la pendiente de la línea que conecta V_1 y V_2 (véase la figura 2-11).

3. Iguale $S_0 = V_2 + G_0[(N - 1)/2]$. Con esto se estima el valor de la serie en el tiempo $t = 0$. Observe que S_0 es el valor asumido por la línea que conecta V_1 y V_2 en $t = 0$ (véase la figura 2-11).

4. *a*) Los factores estacionales iniciales se calculan para cada periodo del cual tenemos datos disponibles y después se promedian para obtener un conjunto de factores estacionales. Los factores estacionales iniciales se obtienen al dividir cada una de las observaciones inicia-

FIGURA 2-11
Inicialización del método de Winters.



les entre el punto a lo largo de la línea que conecta V_1 y V_2. Esto se puede hacer gráficamente o usando la siguiente fórmula

$$c_t = \frac{D_t}{V_i - [(N+1)/2 - j]G_0} \qquad \text{para } -2N+1 \leq t \leq 0,$$

donde $i = 1$ para la primera estación, $i = 2$ para la segunda estación y j es el periodo de la estación. Esto es, $j = 1$ para $t = -2N + 1$ y $t = -N + 1$; $j = 2$ para $t = -2N + 2$ y $t = -N + 2$, y así sucesivamente.

b) Promedie los factores estacionales. Asumiendo exactamente dos estaciones de datos iniciales, obtenemos

$$c_{-N+1} = \frac{c_{-2N+1} + c_{-N+1}}{2}, \ldots, c_0 = \frac{c_{-N} + c_0}{2}.$$

c) Normalicemos los factores estacionales

$$c_j = \left[\frac{c_j}{\sum_{i=0}^{-N+1} c_i}\right] \cdot N \qquad \text{para } -N+1 \leq j \leq 0.$$

Este procedimiento de inicialización que acaba de analizarse es el sugerido por Winters. No es el único medio para inicializar el sistema. Los factores estacionales pueden determinarse por el método de promedios móviles descrito en la primera parte de esta sección. Otra alternativa puede ser ajustar la regresión lineal a los datos de base y usar los valores de la pendiente resultante y la intercepción, como se hizo con el método de Holt, para obtener S_0 y G_0. Los factores estacionales podrán obtenerse al dividir cada observación de demanda en el periodo base por el valor correspondiente de la línea de regresión, promediando periodos iguales, y normalizando. Los valores reales de los estimados iniciales de la intercepción, la pendiente y los factores estacionales serán similares sin importar el esquema de inicialización que se utilice.

Ejemplo *2.8*

Suponga que el conjunto de datos inicial es el mismo que el del ejemplo 2.7, donde se usaron los promedios móviles para encontrar los factores estacionales. Recuerde que tenemos dos estaciones de datos: 10, 20, 26, 17, 12, 23, 30, 22. Entonces

$$V_1 = (10 + 20 + 26 + 17)/4 = 18.25,$$
$$V_2 = (12 + 23 + 30 + 22)/4 = 21.75,$$
$$G_0 = (21.75 - 18.25)/4 = 0.875,$$
$$S_0 = 21.75 + (0.875)(1.5) = 23.06.$$

Los factores estacionales iniciales se calculan de la siguiente manera:

$$c_{-7} = \frac{10}{18.25 - (5/2 - 1)(0.875)} = 0.5904,$$

$$c_{-6} = \frac{20}{18.25 - (5/2 - 1)(0.875)} = 1.123.$$

Los otros factores se calculan de forma similar. Son los siguientes

$$c_{-5} = 1.391, \qquad c_{-4} = 0.869, \qquad c_{-3} = 0.5872,$$
$$c_{-2} = 1.079, \qquad c_{-1} = 1.352, \qquad c_0 = 0.9539.$$

Después se promedian c_{-7} y c_{-3}, c_{-6} y c_{-2}, y así sucesivamente, para obtener los cuatro factores estacionales:

$$c_{-3} = 0.5888, \quad c_{-2} = 1.1010, \quad c_{-1} = 1.3720, \quad c_0 = 0.9115.$$

Por último, se normalizan los factores para asegurar que sumen 4, se obtiene

$$c_{-3} = 0.5900, \quad c_{-2} = 1.1100, \quad c_{-1} = 1.3800, \quad c_0 = 0.9200.$$

Observe lo cerca que concuerdan estos factores con los obtenidos por medio del método de promedios móviles.

Suponga que deseamos pronosticar la demanda del próximo año en el tiempo $t = 0$. La ecuación de pronóstico es

$$F_{t,\,t+\tau} = (S_t + \tau G_t)c_{t+\tau-N},$$

que da como resultado

$$F_{0,1} = (S_0 + G_0)c_{-3} = (23.06 + 0.875)(0.59) = 14.12,$$
$$F_{0,2} = (S_0 + 2G_0)c_{-2} = [23.06 + (2)(0.875)](1.11) = 27.54,$$
$$F_{0,3} = 35.44,$$
$$F_{0,4} = 24.38.$$

Ahora suponga que en el tiempo $t = 1$ se observa una demanda $D_1 = 16$. Ahora es necesario actualizar las ecuaciones. Suponga que $\alpha = .2$, $\beta = .1$ y $\gamma = .1$. Entonces

$$S_1 = \alpha(D_1/c_{-3}) + (1 - \alpha)(S_0 + G_0)$$
$$= (0.2)(16/.59) + (0.8)(23.06 + 0.875) = 24.57,$$
$$G_1 = \beta(S_1 - S_0) + (1 - \beta)G_0$$
$$= (0.1)(24.57 - 23.06) + (0.9)(0.875) = 0.9385,$$
$$c_1 = \gamma(D_1/S_1) + (1 - \gamma)c_{-3}$$
$$= (0.1)(16/24.57) + (0.9)(0.59) = 0.5961.$$

En este punto, se puede volver a normalizar c_{-2}, c_{-1}, c_0, y el nuevo valor de c_1 para que sumen 4. Como la suma es 4.0061, está lo suficientemente cerca (la normalización dará como resultado un redondeo de c_1 a .59 de nuevo).

Al realizar un pronóstico a partir del periodo 1, se obtiene

$$F_{1,2} = (S_1 + G_1)c_{-2} = (24.57 + 0.9385)(1.11) = 28.3144,$$
$$F_{1,3} = (S_1 + 2G_1)c_{-1} = [24.57 + (2)(0.9385)](1.38) = 36.4969,$$

y así sucesivamente.

Ahora suponga que se ha observado un año completo de demanda, dado por $D_1 = 16$, $D_2 = 33$, $D_3 = 34$ y $D_4 = 26$. Cada vez que está disponible una nueva observación, se actualizan los estimados de la intercepción, la pendiente y los factores estacionales. Se obtiene

$$S_2 = 26.35, \quad S_3 = 26.83, \quad S_4 = 27.89,$$
$$G_2 = 1.0227, \quad G_3 = 0.9678, \quad G_4 = 0.9770,$$
$$c_2 = 1.124, \quad c_3 = 1.369, \quad c_4 = 0.9212.$$

Como c_1, c_2, c_3 y c_4 suman 4.01, no es necesaria la normalización. Suponga que estábamos interesados en el pronóstico realizado en el periodo 4 para el periodo 10. La ecuación de pronóstico es

$$F_{t,\,t+\tau} = (S_t + \tau G_t)c_{t+\tau-2N},$$

que da como resultado

$$F_{4,10} = (S_4 + 6G_4)c_2 = [27.89 + 6(0.9770)](1.124) = 37.94.$$

Una consideración importante es elegir constantes de suavizamiento α, β y γ para usarlas en el método de Winters. Los problemas aquí son los mismos que los descritos para el suavizamiento exponencial simple y el método de Holt. Los valores grandes de las constantes resultarán en pronósticos más sensibles pero menos estables. Un método para establecer α, β y γ es experimentar con varios valores de los parámetros que retrospectivamente se ajustan mejor a los pronósticos previos al historial observado de la serie. Como debemos probar muchas combinaciones de las tres constantes, los cálculos son muy tediosos. Aún más, no existe garantía de que los mejores valores de las constantes de suavizamiento basados en datos pasados serán los mejores para pronósticos futuros. La postura más conservadora es garantizar pronósticos estables eligiendo constantes de suavizamiento entre 0.1 y 0.2.

Más problemas para la sección 2.9

35. Considere los datos para el problema 34.
    a) Usando los datos de 1991 y 1992, determine los valores iniciales de la intercepción, la pendiente y los factores estacionales para el método de Winters.
    b) Suponga que la demanda observada para el primer trimestre de 1993 fue 18. Usando $\alpha = .2$, $\beta = .15$ y $\gamma = .10$, actualice los estimados de la serie, la pendiente y los factores estacionales.
    c) ¿Cuáles son los pronósticos realizados al final del primer trimestre de 1993 para los tres trimestres restantes de este mismo año?
36. Suponga que la demanda trimestral observada para 1993 fue 18, 51, 86, 66. Compare la exactitud de los pronósticos obtenidos para los tres últimos trimestres de 1993 en los problemas 34*d*) y 35*c*) calculando tanto la DAM como el ECM.
37. Determine los estimados actualizados de la pendiente, la intercepción y los factores estacionales para el final de 1993 con base en las observaciones dadas en el problema 36. Usando los mismos estimados actualizados, determine los pronósticos que se obtienen mediante el método de Winters para todo 1995, hechos a finales de 1993. Use los valores de las constantes de suavizamiento que aparecen en el problema 35*b*).

2.10 CONSIDERACIONES PRÁCTICAS

Identificación y monitoreo del modelo

La determinación del modelo adecuado depende tanto de las características del historial de observaciones como del contexto en el que se requieren los pronósticos. Cuando los datos históricos están disponibles, deben examinarse con mucho cuidado para determinar si existen patrones obvios, como la tendencia o las fluctuaciones estacionales. Normalmente, estos patrones pueden localizarse graficando los datos. Es posible utilizar pruebas estadísticas, como la significancia de una regresión para verificar si existe una tendencia. La identificación de relaciones más complejas requiere de métodos con mayor sofisticación. La *función de autocorrelación* puede revelar intrincadas relaciones que los métodos de graficación simple no pueden. La metodología Box-Jenkins (1970) se basa en la identificación del modelo apropiado a partir de un examen de la estructura de autocorrelación. Los métodos Box-Jenkins requieren un historial bastante grande (se recomiendan cuando menos 72 datos) y considerar por separado la optimización y la identificación de un modelo para cada una de las series a considerar. Según la literatura, dichos métodos se usan rara vez en las aplicaciones de producción.

Una vez elegido un modelo, deben monitorearse los pronósticos con regularidad para ver si el modelo es adecuado o si ha ocurrido un cambio imprevisto en la serie. Como indicamos,

un método de pronóstico no debe ser sesgado. Esto es, el valor esperado del error de pronóstico debe ser cero. Además de los métodos mencionados en la sección 2.6, un medio de monitorear el sesgo es la señal de seguimiento desarrollada por Trigg (1964). Si utilizamos la notación anterior, podemos decir que e_t es el error observado en el periodo t y $|e_t|$ el valor absoluto del error observado. Los valores suavizados del error y el error absoluto están dados por

$$E_t = \beta e_t + (1 - \beta)E_{t-1},$$
$$M_t = \beta|e_t| + (1 - \beta)M_{t-1}.$$

La señal de seguimiento es la relación

$$T_t = \left|\frac{E_t}{M_t}\right|.$$

Si los pronósticos son insesgados, el error suavizado E_t debe ser menor en comparación con el error absoluto suavizado M_t. Por lo tanto, un valor grande de la señal de seguimiento indica pronósticos sesgados, lo cual sugiere que el modelo de pronóstico es inadecuado. El valor de T_t que señala un sesgo significativo depende de la constante de suavizamiento β. Por ejemplo, Trigg (1964) afirma que un valor de T_t que excede 0.51 indica errores no aleatorios para una β de .1. La señal de seguimiento puede usarse también directamente como una constante de suavizamiento variable. Esto se considera en el problema 48.

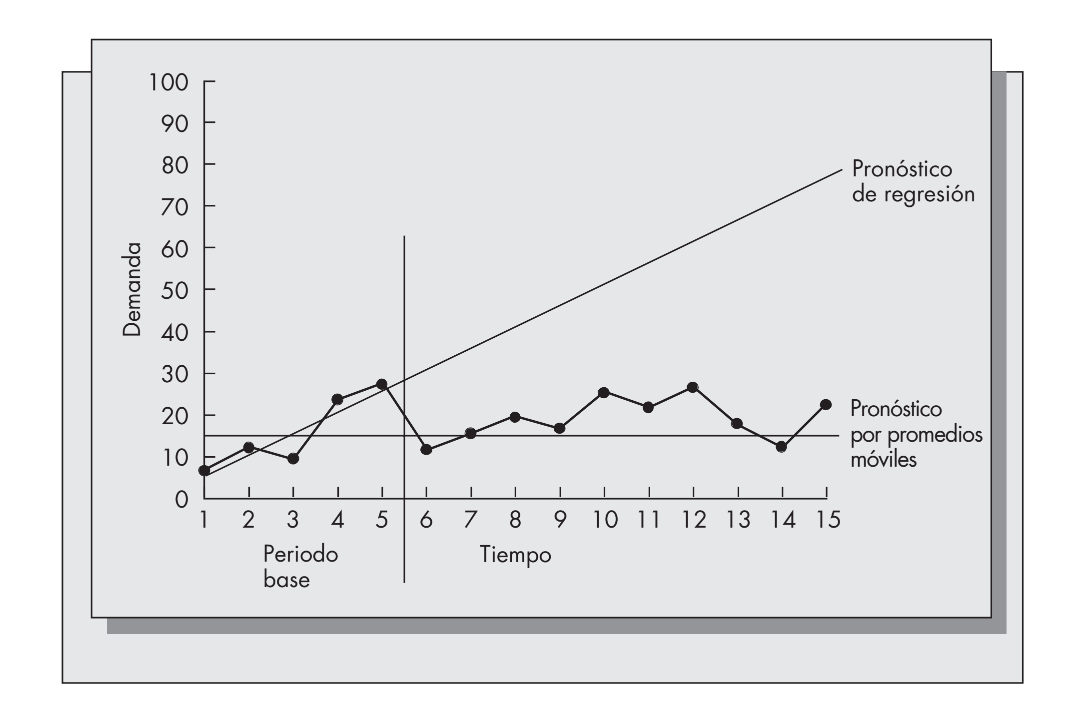
Métodos de series de tiempo simples contra complejos

La literatura sobre pronósticos es voluminosa. En este capítulo hemos tocado solamente algunas técnicas medianamente sencillas. El lector se preguntará si estos métodos realmente funcionan. Los resultados provenientes de la literatura sugieren que los métodos más sencillos son con frecuencia tan exactos como los más sofisticados. Armstrong (1984) revisó 25 años de casos de estudios de pronósticos para averiguar si los métodos sofisticados funcionan mejor. Al comparar los resultados de 39 casos de estudio, descubrió que en 20 casos los métodos sofisticados funcionaron prácticamente de la misma manera que los sencillos, en 11 casos superaron a los sencillos y en 7 resultaron significativamente peores.

Un método de pronóstico más sofisticado es aquel que requiere estimar un número mayor de parámetros a partir de los datos. Pueden surgir problemas cuando estos parámetros se calculan incorrectamente. Para dar una idea de la naturaleza de este problema, consideremos una comparación de promedios móviles sencillos y el análisis de regresión de las siguientes series: 7, 12, 9, 23, 27. Supongamos que estamos interesados en pronosticar, al final del periodo 5, la demanda del periodo 15 (es decir, requerimos $F_{5,\,15}$). El pronóstico por promedios móviles de cinco periodos hecho al final del periodo 5 es 15.6, y esto podrá ser el pronóstico para el periodo 15. El ajuste por mínimos cuadrados de los datos es $\hat{D}_t = 0.3 + 5.1_t$. Al sustituir $t = 15$ obtenemos el pronóstico de regresión de 76.8. En la figura 2-12 vemos la realización de la demanda hasta el periodo 15. Observemos lo que sucedió. La aparente tendencia que existió en los primeros cinco periodos dados fue extrapolada al periodo 15 por la ecuación de regresión. Sin embargo, en realidad no existió una tendencia significativa en este caso en particular. El modelo más complejo dio resultados *significativamente* más pobres para el pronóstico a largo plazo.

Existe cierta evidencia de que el promedio aritmético de los pronósticos obtenido a partir de métodos diferentes resulta más exacto que un método sencillo (véase Makridakis y Winkler, 1983). Quizás esto se debe a que, con frecuencia, un método simple no puede captar la señal subyacente en los datos y los diferentes modelos captan diferentes aspectos de la señal (véase una discusión sobre este fenómeno en Armstrong, 1984.)

FIGURA 2-12
La dificultad con los pronósticos de largo plazo.



¿Qué nos dicen estas observaciones acerca de la aplicación de técnicas de pronóstico para la planeación de producción? En un nivel más alto de planeación, la exactitud del pronóstico es extremadamente importante y los pronósticos de varios pasos adelante juegan un papel integral en la planeación de la fuerza de trabajo y los niveles de producción. Por ello no se aconseja confiar ciegamente en los métodos de series de tiempo en este nivel. En un nivel más bajo dentro del sistema, como puede ser la administración de inventarios para las refacciones, el uso de métodos de series de tiempo simples como los promedios móviles o el suavizamiento exponencial tienen un gran sentido. A nivel de artículos individuales, se requiere de pronósticos de corto plazo para un número más grande de artículos, y el monitoreo de cada artículo resulta impráctico en el mejor de los casos. El riesgo de errores severos se minimiza al utilizar métodos sencillos.

2.11 PANORAMA GENERAL DE TEMAS AVANZADOS EN PRONÓSTICOS

Métodos Box-Jenkins

Este capítulo ha proporcionado una descripción de los métodos de pronóstico más populares para las aplicaciones de planeación de la producción. La mayor parte de la discusión se concentró en los métodos con series de tiempo. Éstos son métodos automáticos que basan los pronósticos en un historial de observaciones. Una técnica avanzada y muy popular con series de tiempo es la denominada método de Box-Jenkins (Box y Jenkins, 1970). Este método explota posibles dependencias entre valores de la serie de periodo a periodo. (Los otros métodos que analizamos en este capítulo suponen que las observaciones sucesivas son independientes.) La consideración de estas dependencias puede mejorar sustancialmente dichos pronósticos.

Por ejemplo, supongamos que una compañía ha instituido un sistema de incentivos de ventas que proporciona un bono al empleado que muestre una mayor mejoría en apartados de un mes a otro. Con un incentivo como ése, a un mes de malas ventas con frecuencia le sigue un mes de buenas ventas. De forma similar, un mes de buenas ventas normalmente estará seguido por un

momento de calma. Esto significa que las ventas en meses consecutivos tienden a correlacionarse negativamente. Tal información puede usarse para mejorar los pronósticos de ventas.

La *autocorrelación* es la correlación (es decir, el grado de dependencia) entre valores de datos observados separados por un número fijo de periodos. En el ejemplo del párrafo anterior podemos decir que la serie tiene una autocorrelación negativa de primer orden. Si los valores altos tienden a ser seguidos por valores altos, la autocorrelación de primer orden es positiva. También podemos medir autocorrelaciones de mayor orden que el primero. Por ejemplo, si los valores bajos tienden a ser seguidos por valores bajos dos periodos después, entonces la serie muestra una autocorrelación positiva de segundo orden. Para usar estos métodos se determinan las autocorrelaciones de las muestras a partir de los datos de un periodo de separación, dos periodos de separación, etcétera.

El modelo Box-Jenkins más general se conoce como el modelo de promedios móviles autorregresivo integrado (ARIMA, *autoregressive integrated moving average*). Los procesos autorregresivos y los procesos de promedios móviles (que no deben confundirse con los promedios móviles sencillos analizados anteriormente en este capítulo) son dos modelos de series de tiempo con autocorrelación. El término "integrado" se refiere a diferenciación. La diferenciación de primer orden significa derivar una nueva serie que es igual a las primeras diferencias de valores sucesivos de la serie original. La diferenciación es un medio para eliminar la tendencia y el crecimiento polinomial.

Los lectores interesados deben consultar tanto a Box y Jenkins (1970) como a Nelson (1973) para obtener información completa. Los modelos ARIMA no son tan sencillos de construir como otros métodos comentados en este capítulo. Desarrollar la forma adecuada de modelo requiere una profunda comprensión de estos procesos. Es necesario emplear técnicas matemáticas sofisticadas para estimar los parámetros del modelo. Aún más, para obtener una estimación razonable de la función de autocorrelación, debemos contar con un historial de observaciones sustancial. Se recomienda contar con cuando menos 72 datos. La recompensa es que bajo las circunstancias correctas, estos métodos pueden funcionar significativamente mejor que los más sencillos.

Simulación como una herramienta de pronóstico

La simulación en computadora es una técnica muy poderosa para abordar problemas complejos. Una simulación en computadora es la descripción de un problema reducido a un programa de computadora. El programa está diseñado para recrear los aspectos clave de la dinámica de una situación real. Cuando un problema es demasiado complejo para modelarse matemáticamente, la simulación es una alternativa muy popular. Al volver a correr el programa bajo diferentes condiciones de arranque y/o diferentes escenarios, podemos, mediante un proceso de prueba y error, descubrir la mejor estrategia para manejar un sistema.

La simulación es una herramienta común en los problemas de planeación de manufactura como los complejos problemas de flujo de materiales en una planta. Se utiliza mucho menos como una herramienta de pronóstico. Compaq Computer, un exitoso productor de computadoras personales con sede en Houston, Texas, ha experimentado con una poderosa herramienta de pronóstico basada en la simulación para ayudar al proceso de introducción de nuevos productos (McWilliams, 1995). El programa recomienda un momento de lanzamiento y un precio óptimo para la introducción de nuevos productos mediante la incorporación de pronósticos de disponibilidad de componentes y cambios en los precios, de la demanda fluctuante para una característica o un precio dado, y el impacto de los modelos rivales.

Con esta herramienta, Compaq decidió retrasar el lanzamiento de varios modelos equipados con Pentium a finales de 1994. Esta estrategia "era contaria a todo lo que la compañía creía". La estrategia básica de Compaq había sido siempre ser un líder en tecnología. Pero su herramienta de pronóstico sugirió que a finales de 1994 los clientes corporativos no estaban totalmente listos para cambiar a las máquinas Pentium. La estrategia probó ser muy redituable para Compaq, que posteriormente registró ganancias récord.

Pronóstico de la demanda en presencia de ventas perdidas

Los detallistas confían mucho en los pronósticos. Los artículos básicos (los que no cambian mucho de una estación a la siguiente, como las camisas de vestir para caballeros) tienen, por lo general, un historial bastante amplio, lo cual favorece el uso de métodos de series de tiempo para pronosticar su demanda. Sin embargo, hay una diferencia importante entre lo que se observa y lo que se quiere pronosticar. El objetivo es pronosticar la *demanda*, pero sólo se observan las *ventas*. ¿Cuál es la diferencia? Suponga el lector que una clienta quiere comprar una blusa de cierta talla y color, y ve que no está en el anaquel. ¿Qué hará? Quizá pedirá atención especial por parte de un vendedor, pero lo más probable es que salga de la tienda y trate de encontrar la prenda en algún otro lugar. A esto se le llama una venta perdida. La dificultad es que la mayoría de los detallistas no tienen forma de rastrear las ventas perdidas. Observan las ventas, pero necesitan estimar la demanda.

Por ejemplo, considere un artículo que se reabastece para tener 10 unidades al principio de cada semana. Suponga que en las últimas 15 semanas la historia de las ventas de ese artículo fue 7, 5, 10, 10, 8, 3, 6, 10, 10, 9, 5, 0, 10, 10, 4. Fíjese en las semanas en las que las ventas fueron 10 unidades. ¿Cuál fue la demanda en ellas? La respuesta es que no se conoce. Sólo se sabe que fue 10, *cuando menos*. Si usted calcula el promedio y la varianza de esta muestra de números, subestimaría la media y la varianza reales para la demanda.

¿Cómo pronosticar la demanda en este caso? En estadística clásica esto se conoce como muestra recortada. Quiere decir que sólo se conocen los valores de la demanda para una parte de la muestra. Para la otra parte sólo conocemos la cota inferior de la demanda. Hay métodos estadísticos especiales que incorporan el recorte y que producen estimados bastante mejores de la media y la varianza poblacionales en este caso. Esos métodos se pueden incrustar en esquemas de pronósticos secuenciales, como el suavizamiento exponencial, para producir pronósticos mucho mejores.

Nahmias (1994) describió el problema de pronosticar en presencia de ventas perdidas cuando la distribución real de la demanda es normal. Comparó el método de máxima verosimilitud para muestras recortadas con un nuevo método, cualquiera de los cuales se podría incorporar en las rutinas de suavizamiento exponencial. Demostró también que ambos métodos daban como resultado pronósticos bastante mejores tanto de la media como de la variación de la demanda.

Para ver lo impactante que puede ser esa diferencia pongamos un caso en el que la demanda real semanal de un producto es una variable aleatoria normal con media 100 y desviación estándar 30. Supongamos que los artículos se surten hasta llegar a 110 unidades al inicio de cada semana. Se usa el suavizamiento exponencial para obtener dos conjuntos de pronósticos: el primero tiene en cuenta las ventas perdidas (incluye el recorte) y el segundo no (no incluye el recorte). Las figuras 2-13 y 2-14 muestran los estimadores de la media y la desviación estándar con y sin censura. Observe el gran sesgo hacia abajo cuando no se toman en cuenta las ventas perdidas en ambos casos. Esto quiere decir que cuando no se tiene en cuenta en forma correcta la diferencia entre las ventas y la demanda, se subestima tanto la media como la varianza de la demanda. Como ambas cosas son datos para determinar los niveles óptimos de reabastecimiento, éstos podrían estarse subestimando mucho.

2.12 VINCULACIÓN DEL PRONÓSTICO Y LA ADMINISTRACIÓN DE INVENTARIOS

En el capítulo 5 trataremos con mayor detalle el control de inventarios bajo situaciones de incertidumbre de demanda. En la práctica la administración de inventarios y el pronóstico de demanda están fuertemente vinculados. El método de pronóstico puede ser cualquiera de los métodos

FIGURA 2-13
Seguimiento de la media en presencia de ventas perdidas.

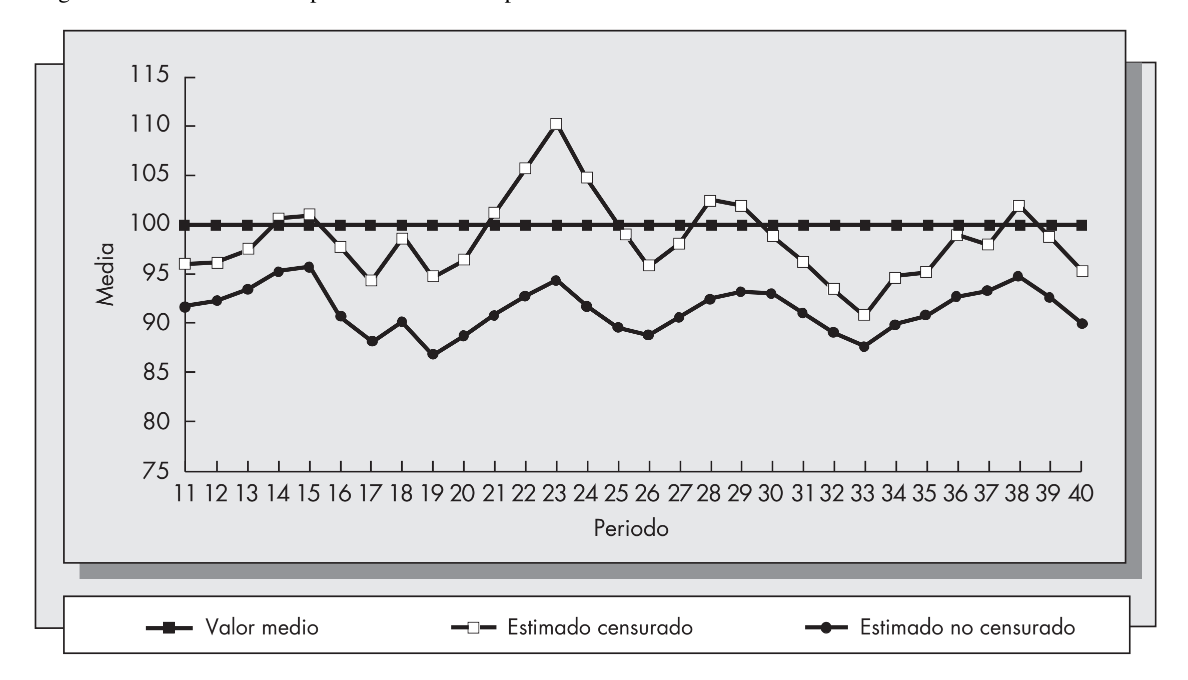
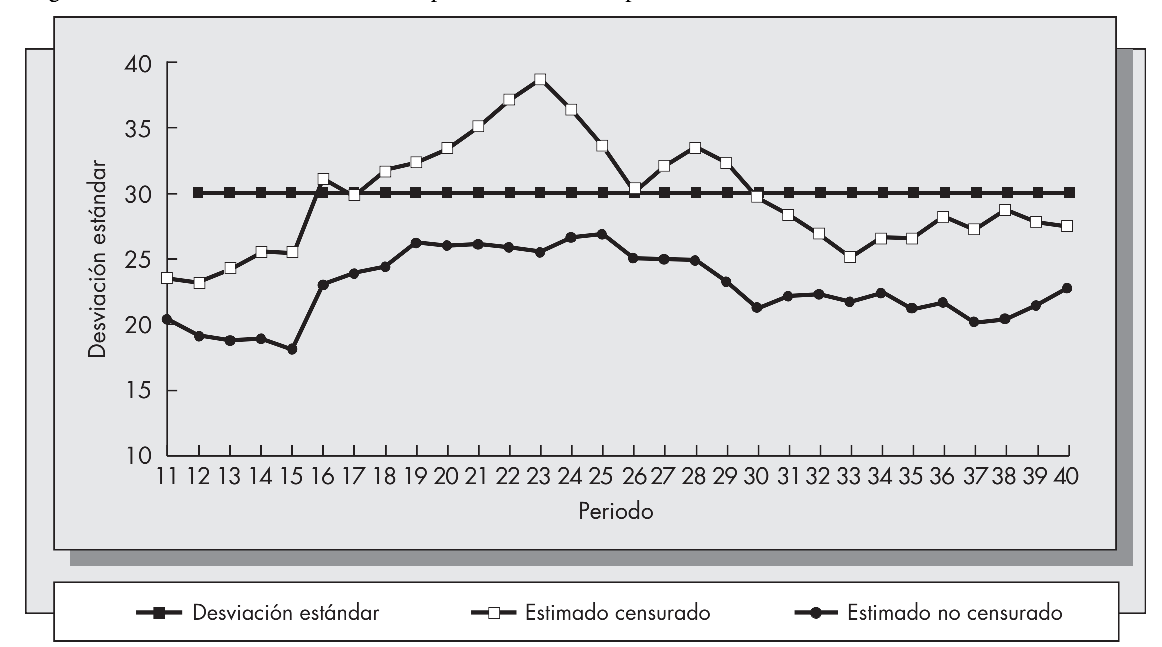


FIGURA 2-14
Seguimiento de la desviación estándar en presencia de ventas perdidas.



comentados en este capítulo. Una de las entradas requeridas para los modelos de control de inventarios es la distribución de la demanda en un periodo, o bien, sobre un tiempo de demora para el reabastecimiento de un pedido. ¿Existe un vínculo entre la distribución de los errores de pronóstico y la distribución de la demanda? La respuesta es sí. La distribución del error de programación juega un papel clave en la correcta aplicación de los modelos de inventarios en la práctica.

El primer asunto es decidir la forma adecuada de distribuir la demanda. La mayoría de los sistemas comerciales suponen que la distribución de la demanda es normal. Eso significa que sólo necesitamos estimar la media μ y la desviación estándar σ para especificar toda la distribución. (Los métodos estadísticos, conocidos como técnicas de bondad de ajuste, pueden aplicarse para probar la exactitud del supuesto de normalidad.) Esté o no vinculado el sistema de inventarios al sistema de pronóstico, debemos contar con un historial de observaciones de la demanda para obtener cálculos estadísticos de la media y la varianza. (Cuando no existe historial de la demanda, el juicio personal debe sustituir a la estimación estadística. En la sección 2.3 se analizan algunos métodos para la agregación de juicios subjetivos.)

Al principio del capítulo 5 comentamos la forma en que debemos estimar la media y la varianza de la demanda directamente de un historial de observaciones de la demanda. En la práctica, sin embargo, no utilizamos generalmente un largo historial de observaciones, porque creemos que la distribución de demanda subyacente no permanece indefinidamente constante. Es por esto que ajustamos la N para los promedios móviles o la α para el suavizamiento exponencial, a fin de balancear la estabilidad y el grado de reacción. Si éste es el caso, ¿cuál es el estimador apropiado de la varianza que se debe utilizar?

En el apéndice 2-A mostramos que para promedios móviles sencillos de orden N, la varianza del error de pronóstico, σ_e^2, está dado por

$$\sigma_e^2 = \sigma^2\left(\frac{N+1}{N}\right)$$

y para el suavizamiento exponencial la varianza del error de pronóstico es

$$\sigma_e^2 = \sigma^2\left(\frac{2}{2-\alpha}\right).$$

Obsérvese que en ambos casos el valor de σ_e^2 *es mayor* que el de σ^2. También obsérvese que cuando N se hace grande y α pequeña, los valores de σ_e y de σ se acercan. Esos casos se presentan al usar el total de las observaciones de la historia de la demanda para elaborar un pronóstico. En el capítulo 5, uno de los datos necesarios para calcular las existencias de seguridad para inventarios es la distribución de la demanda. El problema es que si se cuenta con estimadores de σ_e y de σ, ¿cuál debe usarse como estimador de la desviación estándar para establecer los inventarios de seguridad?

La respuesta obvia es que se debe usar el estimador de σ, porque representa la desviación estándar de la demanda. Sin embargo, la respuesta correcta es que se debería usar σ_e. La razón es que el proceso de pronóstico introduce un error muestral en el proceso de estimación, y ese error se tiene en cuenta en el valor de σ_e. La varianza del error de pronóstico es mayor que la varianza de la demanda, porque el pronóstico sólo se basa en una parte limitada del historial de la demanda.

También puede darse una explicación intuitiva. Si se usa un pronóstico para estimar la media de la demanda, hay que mantener existencias de seguridad para protegerse contra el error en este pronóstico. Por consiguiente, la distribución de los errores de pronóstico es más importante que la de las demandas. Este es un asunto práctico que ha sido fuente de mucha confusión en las publicaciones. El tema se retomará en el capítulo 5, donde se describirá su relevancia en el contexto de los cálculos de existencias de seguridad.

La mayoría de los sistemas de control de inventarios usan el método sugerido por R. G. Brown (1959 y 1962) para estimar el valor de σ_e. (De hecho, parece que Brown fue el primero en reconocer la importancia de la distribución de los errores de pronóstico en las aplicaciones de administración de inventarios.) El método requiere estimar la desviación absoluta media (DAM) de los errores de pronóstico, mediante suavizamiento exponencial. Esto se logra con la ecuación de suavizamiento

$$\text{DAM}_t = \alpha|F_t - D_t| + (1-\alpha)\text{DAM}_{t-1}$$

La DAM se convierte en un estimado de la desviación estándar del error de pronóstico al multiplicarla por 1.25. Esto es, el estimador para σ_e obtenido para el tiempo t es

$$\hat{\sigma}_e = 1.25 \text{ DAM}_t.$$

Se usa un valor pequeño de α, por lo general entre 0.1 y 0.2, para asegurar la estabilidad en el estimador de la DAM. Este método para la estimación de la DAM funciona en cualquiera de los métodos que se describen en este capítulo. A continuación se calculan las existencias de seguridad usando este estimador de σ_e.

Aplicación instantánea

PREDICCIÓN DE RECESIONES ECONÓMICAS

El inicio y el final de los ciclos de negocios son difíciles de predecir. La duración del auge y el ocaso puede variar. Si fuésemos capaces de pronosticar con exactitud el surgimiento de una recesión, los gobiernos y las empresas podrían tomar precauciones para enfrentarlos. Resulta que los economistas han podido identificar un indicador económico importante de las recesiones. Esto es, pueden predecir la probabilidad de que haya una recesión el siguiente año con base en el valor de este indicador. El indicador es la *curva de rendimiento*, que se define como la diferencia entre la tasa de interés en instrumentos de deuda del Tesoro y un bono del Tesoro a tres meses. Estrella y Mishkin (1996)[1] han mostrado que la curva de rendimiento proporciona un excelente pronosticador de la probabilidad de una recesión con cuatro trimestres de anticipación. Usando los datos de 1960 a 1995, los autores han recopilado los siguientes resultados usando un modelo probado:

Probabilidad de recesión (porcentaje)	Dispersión de la curva de rendimiento (en puntos porcentuales)
5	1.21
10	0.76
15	0.46
20	0.22
25	0.02
30	−0.17
40	−0.50
50	−0.82
60	−1.13
70	−1.46
80	−1.85
90	−2.40

[1] Estrella, A. y Mishkin, F. S., “The Yield Curve as a Predictor of U.S. Recessions”, *Current Issues in Economics and Finance*, The Federal Reserve Bank of New York, volumen 2, núm. 7, junio de 1996, pp. 1-6.

Muchos de nosotros probablemente habremos escuchado el término “curva de rendimiento invertida”. Éste se refiere al caso en el que los porcentajes de corto plazo excedieron a los de largo plazo (los valores negativos en la tabla). Como podemos ver, esto es una clara señal de que puede presentarse una recesión en los próximos cuatro trimestres. Como un ejemplo de la forma en que podemos aplicar estos resultados, consideremos el tercer trimestre de 1994. La extensión de la curva de rendimiento se expandió a un promedio de 2.74% y este fue el inicio de uno de los auges más grandes de la historia reciente. En contraste, la extensión de la curva de rendimiento promedió −2.18% en el primer trimestre de 1981, implicando la probabilidad de una recesión en los siguientes cuatro trimestres de 86.5%. Como se predijo, el National Bureau of Economic Research, de hecho, designó el primer trimestre de 1982 como un trimestre de recesión.

2.13 NOTAS HISTÓRICAS Y TEMAS ADICIONALES

Los pronósticos constituyen un área rica para la investigación. No puede exagerarse su importancia en aplicaciones empresariales. Los métodos sencillos de series de tiempo descritos en este capítulo tienen sus raíces en la estadística y la teoría de la probabilidad básicas. El método de suavizamiento exponencial se suele atribuir a R. G. Brown (1959 y 1962), cuando trabajaba como consultor de A. D. Little. Aunque no se fundamenta en la teoría básica, quizás el suavizamiento exponencial sea uno de los métodos más usados de pronóstico en la actualidad. Fue también Brown el primero en reconocer la importancia de la distribución de los errores de pronóstico, y sus implicaciones en la administración de inventarios. Sin embargo, el interés en el uso de métodos estadísticos para pronósticos data de principios del siglo XX o antes. (Véase, por ejemplo, Yule, 1926.)

En este capítulo no se describen los métodos de pronóstico de series de tiempo que usan análisis espectral y espacios de estados. Esos métodos son muy complicados y requieren que en los datos exista bastante estructura. Con frecuencia se basan en el empleo de la función de autocorrelación, y en ese sentido son, conceptualmente, semejantes a los métodos de Box-Jenkins descritos brevemente en la sección 2.11. El trabajo inicial en esta área se debe a Norbert Wiener (1949) y Rudolph Kalman (1960). Sin embargo, se han aplicado poco para pronosticar series económicas de tiempo. La mayor parte de las aplicaciones se ha dado en el área del procesamiento de señales, en ingeniería eléctrica. [Davenport y Root (1958) proporcionan un buen resumen de los conceptos fundamentales de esta área.]

El *filtro de Kalman* es una técnica de suavizamiento exponencial en la que cambia el valor de la constante de suavizamiento conforme pasa el tiempo, y se determina en cierta forma óptima. La idea de ajustar el valor de la constante con base en una medida del desempeño anterior se ha usado en varias formas *ad hoc*. Un procedimiento típico es el que sugieren Trigg y Leach (1967), que requiere el cálculo de la señal de seguimiento. Esta señal se usa como el valor de la constante de suavizamiento para el siguiente pronóstico. La idea es que cuando la señal de seguimiento es grande, indica que la serie de tiempo ha sufrido un desplazamiento; un valor mayor de la constante debe responder más a un desplazamiento brusco en la señal básica. También se han sugerido otros métodos. En forma intencional ha quedado inconcluso un análisis explícito de los métodos adaptativos de rapidez de respuesta por una razón: hay pocas pruebas de que funcionen en el contexto de la predicción de series económicas de tiempo. La mayoría de los estudios que comparan la eficacia de distintos métodos de pronóstico para muchas series distintas no indican ventaja alguna para los modelos adaptativos de rapidez de respuesta. (Véase, por ejemplo, Armstrong, 1984.) El método de Trigg y Leach se describe con más detalle en el problema 48.

Un buen punto de partida para lecturas posteriores es el texto de Makridakis, Wheelwright y McGhee (1983), que contiene una exposición más detallada de los métodos de series de tiempo, de un nivel semejante al que aquí se presenta. Las aplicaciones en las empresas se describen en Wilson y Keating (1990).

2.14 Resumen

En este capítulo se presentó una introducción a varias de las técnicas más usadas para pronosticar series de tiempo, y también una breve descripción de otros métodos, incluyendo el método Delphi y los modelos causales. Un pronóstico de *promedios móviles* se obtiene calculando el promedio aritmético de las N observaciones más recientes de la demanda. Un pronóstico de *suavizamiento exponencial* se determina calculando el promedio ponderado de la observación actual de la demanda y del pronóstico más reciente de esa demanda. El peso que se aplica a la

observación actual es α, y el que se aplica al último pronóstico (esto es, a las observaciones en el pasado) es $1 - \alpha$. Cuando N es pequeña y α es grande, se producen pronósticos cambiantes, y cuando N es grande y α es pequeña, se producen pronósticos estables. Aunque los dos métodos tienen propiedades parecidas, por lo general se prefiere el suavizamiento exponencial, porque sólo requiere guardar el pronóstico anterior, mientras que para los promedios móviles se requiere guardar las últimas N observaciones de la demanda.

Cuando hay una tendencia en la serie, tanto los promedios móviles como el suavizamiento exponencial se retrasan respecto a la tendencia. Describimos dos técnicas de series de tiempo que buscan seguir la tendencia. Una es el *análisis de regresión*, que emplea los mínimos cuadrados para ajustar una recta a los datos, y otra es el *método de Holt*, una especie de suavizamiento exponencial doble. El método de Holt tiene la ventaja de facilitar la actualización de los pronósticos a medida que se dispone de nuevas observaciones de demanda.

También se describen técnicas para series estacionales. Se emplea la *descomposición clásica* con el fin de indicar cómo se usan los promedios móviles simples para estimar los factores estacionales y obtener la serie desestacionalizada cuando hay una tendencia, y se indica cómo pueden estimarse con rapidez los factores estacionales cuando no hay tendencia. La extensión del método de Holt para manejar los problemas con estacionalidad, llamada *método de Winters*, es una técnica de suavizamiento exponencial triple.

La sección 2.11 describe una perspectiva general de los métodos avanzados, más allá del propósito de este texto. La sección final aborda la relación entre los pronósticos y el control de inventarios. El punto clave, que se desarrollará en el capítulo 5, es que la desviación estándar del error de pronóstico es la medida adecuada de la variación para calcular las existencias de seguridad.

Problemas adicionales sobre pronósticos

38. John Kittle, agente independiente de seguros, usa un promedio móvil de cinco años para pronosticar la cantidad de reclamaciones por año para una de las grandes aseguradoras en las que trabaja como vendedor. Acaba de descubrir que un empleado suyo anotó erróneamente la cantidad de reclamaciones de hace cuatro años como 1 400, en lugar de 1 200.

    a) ¿Qué ajuste debe hacer el señor Kittle en el pronóstico del año próximo, para tener en cuenta el valor corregido de la cantidad de reclamaciones hace cuatro años?

    b) Suponga que el señor Kittle usó suavizamiento exponencial doble, con $\alpha = .2$, en lugar de promedios móviles, para determinar su pronóstico. ¿Qué ajuste se requiere ahora para el pronóstico del año próximo? (Observe que se necesita conocer el valor del pronóstico del año próximo para poder resolver este problema.)

39. En la sección 2.12 se describió un método para estimar la desviación absoluta media en donde se recalcula cada vez que se observa una nueva demanda, aplicando la siguiente fórmula:

    $$\text{DAM}_t = \alpha|e_t| + (1 - \alpha)\,\text{DAM}_{t-1}.$$

    Considere los pronósticos de un paso adelante para las fallas de motores de aeronaves para los trimestres 2 a 8, obtenidos en el ejemplo 2.3. Suponga que el valor inicial de la DAM = 50 en el periodo 1. Con el mismo valor de α, ¿qué valores de DAM dan como resultado este método para los periodos 2 a 8? Describa las ventajas y desventajas de este método, comparado con el cálculo directo de la DAM.

40. Herman Hahn trata de establecer un sistema integrado de pronóstico y control de inventarios para su ferretería, Hahn's Hardware. Cuando Herman indica que las lámparas exterio-

res son artículo estacional en su computadora, el programa le solicita teclear los factores estacionales por trimestre.

Desafortunadamente, Herman no ha llevado un registro histórico de datos, pero estima que la demanda de tales lámparas durante el primer trimestre es aproximadamente 30% menor que el promedio, para el segundo trimestre es 20% menor que el promedio, para el tercer trimestre es más o menos igual al promedio, y para el cuarto trimestre la demanda sube al 50% sobre el promedio. ¿Qué debe teclear como factores estacionales?

41. Irwin Richards, editor de textos de negocios, publica en los campos de administración, mercadotecnia, contabilidad, producción, finanzas y economía, entre otros. Al director de la compañía le interesa obtener una medida relativa de los tamaños de libro en los distintos campos. Durante los últimos tres años las cantidades promedio de páginas por libro publicado fueron

	Año 1	**Año 2**	**Año 3**
Administración	835	956	774
Mercadotecnia	620	540	575
Contabilidad	440	490	525
Producción	695	680	624
Finanzas	380	425	410
Economía	1 220	1 040	1 312

Utilice el método rápido descrito en la sección 2.9 para calcular los factores multiplicativos para cada campo, que den el porcentaje sobre o debajo de la cantidad promedio de páginas.



42. Durante un periodo de dos años, la Topper Company vendió las siguientes cantidades de podadoras de pasto:

Mes:	1	2	3	4	5	6	7	8	9	10	11	12
Ventas:	238	220	195	245	345	380	270	220	280	120	110	85
Mes:	13	14	15	16	17	18	19	20	21	22	23	24
Ventas:	135	145	185	219	240	420	520	410	380	320	290	240

a) En la columna A introduzca los números 1 a 24 que representan los meses, y en la B las ventas mensuales. Calcule los pronósticos a tres meses, con promedios móviles, y anótelos en la tercera columna. Asegúrese de alinear su pronóstico con el periodo para el que pronostica. (El promedio de ventas en los meses 1, 2 y 3 se debe poner en el renglón 4, el de las ventas en los meses 2, 3 y 4 en el renglón 5, y así sucesivamente.) En la cuarta columna calcule el error de pronóstico para cada mes en el que haya obtenido un pronóstico.

b) En las columnas 5, 6 y 7 calcule el error absoluto, el error al cuadrado y el error porcentual absoluto. Con estos resultados, determine la DAM, el ECM y el EPAM de los pronósticos PM(3) para los meses 4 a 24.

c) Repita los incisos *a*) y *b*) con promedios móviles de seis meses. (Esos cálculos deben aparecer en las columnas 7 a 11.) ¿Cuál de los métodos, PM(3) o PM(6), fue más exacto para estos datos?



43. Repita los cálculos en el problema 42 usando suavizamiento exponencial simple, y considere que la constante de suavizamiento α es una variable; esto es, la constante debe tener

una dirección de celda. Experimentando con distintos valores de α determine el valor que parece minimizar

a) La DAM.

b) El ECM.

c) El EPAM.

Suponga que el pronóstico para el mes 1 es 225.

44. Baby It's You, fabricante de alimentos para bebé, ha determinado que hay una gran correlación entre las ventas agregadas de la empresa (en 100 000 dólares) y la cantidad de nacimientos a nivel nacional del año anterior. Suponga que las ventas y el número de nacimientos, durante los ocho años anteriores, fueron

	Año							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Ventas (en $100 000)	6.1	6.4	8.3	8.8	5.1	9.2	7.3	12.5
Nacimientos (en millones)	2.9	3.4	3.5	3.1	3.8	2.8	4.2	3.7

a) Suponiendo que los nacimientos en el país representan la variable independiente, y las ventas la dependiente, determine una ecuación de regresión para predecir las ventas con base en los nacimientos. Use los años 2 a 8 como base. (Sugerencia: necesitará las fórmulas generales de regresión que aparecen en el apéndice 2-B para resolver este problema.)

b) Supongamos que se pronostica que habrá 3.3 millones de nacimientos en el año 9. ¿Qué pronóstico para ventas obtendrá en el año 10 usando los resultados del inciso *a*)?

c) Suponga que se usa suavizamiento exponencial simple, con $\alpha = .15$, para predecir la cantidad de nacimientos. Use el promedio de los años 1 a 4 como su pronóstico inicial para el periodo 5, y determine un pronóstico suavizado exponencialmente para los nacimientos en el año 9, en el país.

d) Combine los resultados en *a*), *b*) y *c*) para obtener un pronóstico de la suma de las ventas agregadas totales en los años 9 y 10.

45. Hy y Murray desean establecer una fuente de sodas en Shoreline Park, sitio que se describe en el problema 28. Después de seis meses de funcionamiento, las ventas observadas de helados (en dólares) y la cantidad de asistentes al parque fueron

	Mes					
	1	**2**	**3**	**4**	**5**	**6**
Ventas de helados	325	335	172	645	770	950
Visitantes al parque	880	976	440	1 823	1 885	2 436

a) Determine una ecuación de regresión que considere las ventas de helado como la variable dependiente y el tiempo como la variable independiente. Con base en esta ecuación de regresión, ¿cuánto valdrán las ventas de helados dentro de dos años (mes 30)? ¿Cuánta confianza tiene en este pronóstico? Explique su respuesta.

b) Determine una ecuación de regresión que considere las ventas de helado como la variable dependiente, y los asistentes al parque como la variable independiente. (Sugerencia: Necesitará las ecuaciones generales de regresión del apéndice 2-B para resolver este punto.)

c) Suponga que la asistencia al parque seguirá una curva logística (véase la figura 2-15). El departamento de parques espera que la asistencia sea de 6 000 personas al mes como máximo. Grafique los datos de asistentes al parque, por mes, y trace "a simple vista" una curva logística que se ajuste a los datos, usando 6 000 como el valor máximo. Con base en su curva y en la ecuación de regresión determinada en el inciso *b*), pronostique las ventas de helado para los meses 12 a 18.

46. Un método sugerido para determinar el valor "correcto" de la constante de suavizamiento α es determinar retrospectivamente el valor de α que dé como resultado el mínimo error de pronóstico para algún conjunto de datos históricos. Comente qué tan adecuado es este método y cuáles son algunos de los problemas potenciales que pudieran surgir.

47. Lakeroad, fabricante de discos duros para computadoras personales, fue fundada en 1981 y sus ventas han sido las siguientes:

Año	Cantidad vendida (en miles de unidades)	Año	Cantidad vendida (en miles de unidades)
1981	0.2	1985	34.5
1982	4.3	1986	68.2
1983	8.8	1987	85.0
1984	18.6	1988	58.0

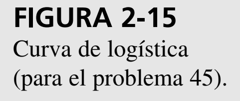
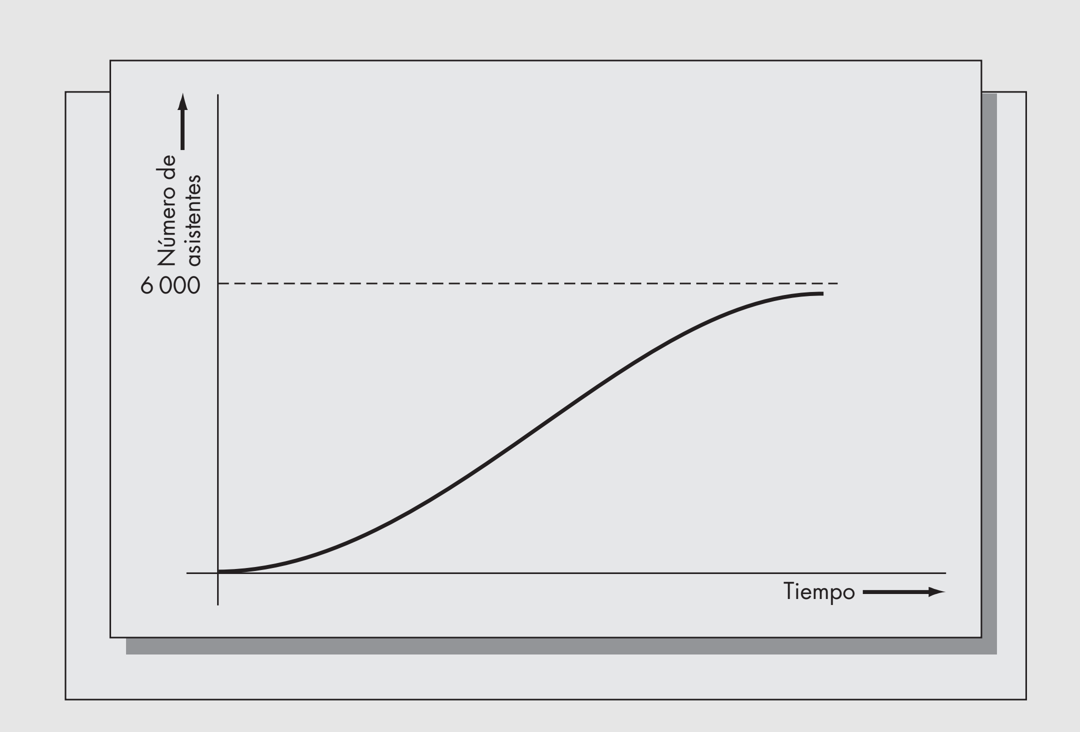


FIGURA 2-15
Curva de logística (para el problema 45).



a) Suponga que en la empresa se usa el método de Holt para pronosticar las ventas. Considere que $S_0 = 0$ y que $G_0 = 8$. Use $\alpha = .2$ y $\beta = .2$ para determinar pronósticos de un paso adelante desde 1982 hasta 1989, y calcule la desviación absoluta media y el error cuadrático medio de los pronósticos durante ese periodo. ¿Cuál es el pronóstico de ventas para el año 2000, formulado al terminar 1988? Con base en los resultados de 1988, ¿por qué podría ser muy impreciso este pronóstico?

b) Haga pruebas con varios valores de α y β y determine los valores de las constantes de suavizamiento que parezcan producir los pronósticos más exactos.

48. Trigg y Leach (1967) sugieren el siguiente método de suavizamiento exponencial adaptativo a la rapidez de respuesta. Junto con el suavizamiento de la serie original, suavice también el error e_t y el error absoluto $|e_t|$ de acuerdo con las ecuaciones siguientes:

$$E_t = \beta e_t + (1 - \beta)E_{t-1},$$

$$M_t = \beta|e_t| + (1 - \beta)M_{t-1}$$

y defina la constante de suavizamiento que se usará para pronosticar la serie en el periodo t como

$$\alpha_t = \left|\frac{E_t}{M_t}\right|.$$

El pronóstico hecho en el periodo t, para el periodo $t + 1$, se obtiene con la ecuación normal de suavizamiento exponencial, usando α_t como la constante de suavizamiento. Esto es,

$$F_{t+1} = \alpha_t D_t + (1 - \alpha_t)F_t.$$

El fundamento del método es que cuando la magnitud de E_t se acerca a M_t, es posible que los pronósticos estén sesgados. En ese caso los resultados obtenidos con un mayor valor de la constante de suavizamiento hacen que el suavizamiento exponencial responda más rápido a los cambios repentinos en la serie.

a) Aplique el método de Trigg y Leach a los datos del problema 22. Use las desviaciones absolutas medias para comparar la exactitud de esos pronósticos con los obtenidos en el problema 22, con suavizamiento exponencial simple. Suponga que $E_1 = e_1$ (el error observado en enero) y que $M_1 = |e_1|$. Use $\beta = .1$.

b) ¿Para qué tipos de series de tiempo dará pronósticos más precisos el método de Trigg-Leach de rapidez de respuesta adaptativa, y bajo qué circunstancias dará pronósticos menos precisos? Comente la conveniencia de usar ese método en casos en los que los pronósticos no se vigilan o monitorean estrechamente.

49. El propietario de una pequeña cervecería aplica el método de Winters para pronosticar sus ventas trimestrales de cerveza. Ha venido usando constantes de suavizamiento $\alpha = .2$, $\beta = .2$ y $\gamma = .2$. Hasta ahora ha obtenido los siguientes valores de pendiente, ordenada al origen y factores estacionales: $S_{10} = 120$, $G_{10} = 14$, $c_{10} = 1.2$, $c_9 = 1.1$, $c_8 = .8$ y $c_7 = .9$.

a) Determine el pronóstico de las ventas de cerveza en el trimestre 11.

b) Suponga que las ventas reales son 128 en el trimestre 11. Calcule S_{11} y G_{11}, y determine los valores actualizados de los factores estacionales. También determine el pronóstico que se hace al final del trimestre 11 para el trimestre 13.



50. El producto nacional bruto (PNB) de Estados Unidos en millardos de dólares durante el periodo de 1964 a 1984 fue el siguiente:

Año	PNB	Año	PNB
1964	649.8	1975	1 598.4
1965	705.1	1976	1 782.8
1966	772.0	1977	1 990.5
1967	816.4	1978	2 249.7
1968	892.7	1979	2 508.2
1969	963.9	1980	2 732.0
1970	1 015.5	1981	3 052.6
1971	1 102.7	1982	3 166.0
1972	1 212.8	1983	3 401.6
1973	1 359.3	1984	3 774.7
1974	1 472.8		

Fuente: *Informe económico del Presidente de Estados Unidos*, febrero de 1986.

a) Aplique el método de Holt para predecir el PNB. Determine un ajuste de los datos entre 1964 y 1974 mediante regresión para estimar los valores iniciales de la pendiente y la intercepción. (Sugerencia: Si calcula la regresión a mano, transforme los años restando 1963 de cada valor, para hacer menos tediosos los cálculos.) Con el método de Holt determine pronósticos para 1975 a 1984. Suponga que $\alpha = .2$ y $\beta = .1$. Calcule la DAM y el ECM de los pronósticos a un paso adelante, para el periodo de 1975 a 1984.

b) Calcule el aumento porcentual del PNB de 1964 a 1984 y grafique la serie resultante. Use un promedio móvil de seis meses y suavizamiento exponencial simple con $\alpha = .2$ para obtener los pronósticos a un paso delante de 1975 a 1984. (Use el promedio aritmético de las observaciones de 1964 a 1974 para inicializar el suavizamiento exponencial.)

En ambos casos (es decir, pronósticos con promedios móviles y con suavizamiento exponencial), convierta sus pronósticos de aumento porcentual para el año siguiente en pronóstico del PNB mismo, y calcule la desviación absoluta media y el error cuadrático medio de cada pronóstico. Compare la precisión de esos métodos con el del inciso *a*).

c) Discuta el problema de predecir el PNB. ¿Qué métodos, además de los que se usaron en los incisos *a*) y *b*), podrían producir predicciones todavía mejores para esta serie?

Apéndice **2-A**

Errores de pronóstico para promedios móviles y suavizamiento exponencial

El error de pronóstico e_t es la diferencia entre el pronóstico para el periodo t y la demanda real para ese periodo. En esta sección deduciremos la distribución del error de pronóstico, tanto para promedios móviles como para suavizamiento exponencial.

Se supone que la demanda se genera con el proceso

$$D_t = \mu + \epsilon_t,$$

donde ϵ_t es normal con media cero y varianza σ^2.

CASO 1. PROMEDIOS MÓVILES

Considere el caso en el que los pronósticos se generan con promedios móviles. Entonces el error de pronóstico es $e_t = F_t - D_t$, donde F_t está dada por

$$F_t = \frac{1}{N} \sum_{i=t-N}^{t-1} D_i.$$

Se deduce que

$$E(F_t - D_t) = (1/N) \sum_{i=t-N}^{t-1} E(D_i) - E(D_t) = (1/N)(N\mu) - \mu = 0.$$

Esto demuestra que cuando la demanda es estacionaria, los pronósticos con promedios móviles no están sesgados.

También,

$$\begin{aligned} \text{Var}(F_t - D_t) &= \text{Var}(F_t) + \text{Var}(D_t) \\ &= (1/N^2) \sum_{i=t-N}^{t-1} \text{Var}(D_i) + \text{Var}(D_t) \\ &= (1/N^2)(N\sigma^2) + \sigma^2 \\ &= \sigma^2(1 + 1/N) = \sigma^2[(N + 1)/N]. \end{aligned}$$

Por lo tanto, la desviación estándar del error de pronóstico, σ_e, es

$$\sigma_e = \sigma\sqrt{\frac{N + 1}{N}}.$$

Esta es la desviación estándar del error de pronóstico para promedios móviles simples, en función de la desviación estándar de cada observación.

Habiendo deducido la media y la varianza del error de pronóstico, todavía falta especificar la forma de la distribución del error de pronóstico. Del supuesto, los valores de D_t forman una sucesión de variables aleatorias normales independientes, idénticamente distribuidas. Como F_t es una combinación lineal de $D_{t-1}, D_{t-2}, \ldots D_{t-N}$, entonces F_t tiene distribución normal y es independiente de D_t. En consecuencia, e_t también es normal. Así, la distribución de e_t se especifica por completo mediante su media y su varianza.

Como el valor esperado del error de pronóstico es cero, se dice que el método es insesgado. Observe que esto es resultado del supuesto que menciona que el proceso de demanda es estacionario. Veamos la variación del error de pronóstico. El valor de N que minimiza a σ_e es $N = +\infty$. Esto quiere decir que la variación se minimiza si el pronóstico es el promedio de todos los datos pasados. Sin embargo, nuestra intuición nos dice que podemos mejorar si usamos datos más recientes para establecer el pronóstico. Surge entonces una discrepancia, porque en realidad no creemos en nuestro supuesto de que el proceso de demanda es estacionario durante todo el tiempo. Un valor menor de N permitiría que el método del promedio móvil reaccionara con más rapidez a cambios imprevistos del proceso de demanda.

CASO 2. SUAVIZAMIENTO EXPONENCIAL

Ahora considere el caso en el que los pronósticos se generan con suavizamiento exponencial. Este caso, F_t se puede representar con la suma ponderada infinita de valores pasados de demanda.

$$F_t = \alpha D_{t-1} + \alpha(1-\alpha)D_{t-2} + \alpha(1-\alpha)^2 D_{t-3} + \cdots,$$
$$E(F_t) = \mu[\alpha + \alpha(1-\alpha) + \alpha(1-\alpha)^2 + \cdots] = \mu.$$

Observe que esto quiere decir que $E(e_t) = 0$, de modo que el suavizamiento exponencial y los promedios móviles son métodos insesgados de pronóstico cuando el proceso subyacente de demanda es una constante más un término aleatorio.

$$\begin{aligned}\text{Var}(F_t) &= \alpha^2\sigma^2 + (1-\alpha)^2\alpha^2\sigma^2 + \cdots \\ &= \sigma^2\alpha^2 \sum_{n=0}^{\infty} (1-\alpha)^{2n}.\end{aligned}$$

Se puede demostrar que

$$\sum_{n=0}^{\infty} (1-\alpha)^{2n} = \frac{1}{1-(1-\alpha)^2}$$

así que

$$\text{Var}(F_t) = \frac{\sigma^2\alpha^2}{1-(1-\alpha)^2} = \frac{\sigma^2\alpha}{2-\alpha}.$$

Como

$$\text{Var}(e_t) = \text{Var}(F_t) + \text{Var}(D_t),$$
$$\text{Var}(e_t) = \sigma^2[\alpha/(2-\alpha) + 1] = \sigma^2[2/(2-\alpha)],$$

o

$$\sigma_e = \sigma\sqrt{\frac{2}{2-\alpha}}.$$

Esta es la desviación estándar del error de pronóstico para suavizamiento exponencial simple, en función de la desviación estándar de cada observación. La distribución del error de pronóstico para suavizamiento exponencial es normal, por razones que son esencialmente iguales a las mencionadas para los promedios móviles.

Observe que si igualamos las varianzas del error de pronóstico para suavizamiento exponencial y para promedios móviles, obtenemos

$$2/(2-\alpha) = (N+1)/N$$

esto es que $\alpha = 2/(N + 1)$, exactamente el mismo resultado que obtuvimos al igualar la edad promedio de los datos para los dos métodos.

Apéndice **2-B**

Deducción de las ecuaciones para la pendiente y la intercepción en el análisis de regresión

En este apéndice deduciremos las ecuaciones para determinar los valores óptimos de a y b en el modelo de regresión. Supongamos que los datos son $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, y que el modelo de regresión que debe ajustarse es $Y = a + bX$. Se define

$$g(a, b) = \sum_{i=1}^{n} [y_i - (a + bx_i)]^2.$$

La función $g(a, b)$ se interpreta como la suma de los cuadrados de las distancias de la recta $a + bx$ a los datos y_i. El objeto del análisis es elegir a y b que minimicen a $g(a, b)$. Esto se logra cuando

$$\frac{\partial g}{\partial a} = \frac{\partial g}{\partial b} = 0.$$

Esto es,

$$\frac{\partial g}{\partial a} = -\sum_{i=1}^{n} 2[y_i - (a + bx_i)] = 0,$$

$$\frac{\partial g}{\partial b} = -\sum_{i=1}^{n} 2x_i[y_i - (a + bx_i)] = 0,$$

que da como resultado dos ecuaciones

$$an + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \tag{1}$$

$$a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i\, y_i. \tag{2}$$

Estas son dos ecuaciones lineales con dos incógnitas a y b. Al multiplicar la ecuación (1) por Σx_i y la ecuación (2) por n se obtiene

$$an\sum_{i=1}^{n} x_i + b\left(\sum_{i=1}^{n} x_i\right)^2 = \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right), \tag{3}$$

$$an\sum_{i=1}^{n} x_i + bn\sum_{i=1}^{n} x_i^2 = n\sum_{i=1}^{n} x_i y_i. \tag{4}$$

Se resta la ecuación (3) de la ecuación (4) y el resultado es

$$b\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right] = n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right). \tag{5}$$

Se hacen las siguientes definiciones $S_{xy} = n\,\Sigma\, x_i y_i - (\Sigma\, x_i)(\Sigma\, y_i)$ y $S_{xx} = n\Sigma\, x_i^2 - (\Sigma x_i)^2$. Con ellas, la ecuación (5) se puede escribir $bS_{xx} = S_{xy}$, de donde se obtiene

$$b = \frac{S_{xy}}{S_{xx}}. \tag{6}$$

De la ecuación (1) tenemos,

$$an = \sum_{i=1}^{n} y_i - b\sum_{i=1}^{n} x_i$$

o

$$a = \bar{y} - b\bar{x}, \tag{7}$$

donde $\bar{y} = (1/n)\Sigma\, y_i$ y $\bar{x} = (1/n)\Sigma\, x_i$.

Estas fórmulas se pueden adaptar al problema de pronosticar cuando se supone que la variable independiente es el tiempo. En ese caso, los datos tienen la forma $(1, D_1)$, $(2, D_2)$, . . . , (n, D_n), y la ecuación de pronóstico tiene la forma $\hat{D}_t = a + bt$. Las diversas fórmulas se pueden simplificar como sigue:

$$\sum x_i = 1 + 2 + 3 + \cdots + n = \frac{n(n + 1)}{2},$$

$$\sum x_i^2 = 1 + 4 + 9 + \cdots + n^2 = \frac{n(n + 1)(2n + 1)}{6}.$$

Por consiguiente, podemos escribir

$$S_{xy} = n \sum_{i=1}^{n} iD_i - n(n+1)/2 \sum_{i=1}^{n} D_i,$$

$$S_{xx} = \frac{n^2(n+1)(2n+1)}{6} - \frac{n^2(n+1)^2}{4}.$$

$$b = \frac{S_{xy}}{S_{xx}}$$

$$a = \overline{D} - \frac{b(n+1)}{2}$$

Apéndice 2-C

Glosario de notaciones para el capítulo 2

a = Estimado de la intercepción, en el análisis de regresión.

α = Constante del suavizamiento exponencial simple. También una de las constantes de suavizamiento usadas en el método de Holt, o una de las constantes de suavizamiento usadas en el método de Winters.

b = Estimado de la pendiente en el análisis de regresión.

β = Segunda constante de suavizamiento usada en el método de Holt o en el de Winters.

c_t = Factor estacional para el t-ésimo periodo de una estación.

γ = Tercera constante de suavizamiento usada en el método de Winters.

D_t = Demanda en el periodo t. Se refiere a la serie cuyos valores se pronosticarán.

e_t = $F_t - D_t$ = Error observado de pronóstico en el periodo t.

ϵ_t = Variable aleatoria que representa el componente aleatorio de la demanda.

F_t = Pronóstico a un paso adelante hecho en el periodo $t - 1$, para la demanda en el periodo t.

$F_{t,\,t+\tau}$ = Pronóstico τ pasos adelante hecho en el periodo $t - 1$, para la demanda en el periodo $t + \tau$.

G_t = Valor suavizado de la pendiente en los métodos de Holt y de Winters.

μ = Media del proceso de demanda.

DAM = Desviación absoluta media = $(1/n) \sum_{i=1}^{n} |e_i|$.

EPAM = Error porcentual absoluto medio = $(1/n) \sum_{i=1}^{n} |e_i/D_i| \times 100$.

ECM = Error cuadrático medio = $(1/n) \sum_{i=1}^{n} e_i^2$.

S_t = Valor suavizado de la serie (intercepción) en los métodos de Holt y de Winters.

σ^2 = Varianza del proceso de demanda.

T_t = Valor de la señal de seguimiento en el periodo t (véase el problema 48).

Bibliografía

Armstrong, J. S. “Forecasting by Extrapolation: Conclusions from Twenty-Five Years of Research”. *Interfaces* 14 (1984), pp. 52-66.

Box, G. E. P. y G. M. Jenkins. *Time Series Analysis*, *Forecasting*, *and Control*. San Francisco: Holden Day, 1970.

Brown, R. G. *Statistical Forecasting for Inventory Control*. Nueva York: McGraw-Hill, 1959.

Brown, R. G. *Smoothing*, *Forecasting*, *and Prediction of Discrete Time Series*. Englewood Cliffs, NJ: Prentice Hall, 1962.

Chambers, J. C., S. K. Mullick y D. D. Smith. "How to Choose the Right Forecasting Technique". *Harvard Business Review* 65 (1971), pp. 45-74.

Davenport, W. B. y W. L. Root. *An Introduction to the Theory of Random Signals and Noise*. Nueva York: McGraw-Hill, 1958.

Draper, N. R. y H. Smith. *Applied Regression Analysis*. Nueva York: John Wiley & Sons, 1968.

Fisher, M. L., J. H. Hammond, W. R. Obermeyer y A. Raman, "Making Supply Meet Demand in an Uncertain World". *Harvard Business Review*, mayo-junio 1994, pp. 221-240.

Gross, C. W. y R. T. Peterson. *Business Forecasting*. 2a. ed. Nueva York: John Wiley & Sons, 1983.

Helmer, O. y N. Rescher. "On the Epistemology of the Inexact Sciences". *Management Science* 6(1959), p. 25-52.

Holt, C. C. "Forecasting Seasonal and Trends by Exponentially Weighted Moving Averages". Office of Naval Research Memorandum, núm. 52, 1957.

Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problems". *Journal of Basic Engineering*, Ser. D 82, 1960, pp. 35-44.

Makridakis, S., S. C. Wheelwright y V. E. McGhee. *Forecasting: Methods and Applications*. 2a. ed. Nueva York: John Wiley & Sons, 1983.

Makridakis, S. y R. L. Winkler. "Averages of Forecasts". *Management Science* 29(1983), pp. 987-996.

McWilliams, G. "At Compaq, a Desktop Crystal Ball". *Business Week*, 20 de marzo de 1995, p. 96-97.

Montgomery, D. C. y L. A. Johnson. *Forecasting and Time Series Analysis*. Nueva York: McGraw-Hill, 1976.

Muth, J. F. "Optimal Properties of Exponentially Weighted Forecasts of Time Series with Permanent and Transitory Components". *Journal of the American Statistical Association* 55(1960), pp. 299-306.

Nahmias, S. "Demand Estimation in Lost Sales Inventory Systems". *Naval Research Logistics* 41(1994), pp. 739-757.

Nelson, C. R. *Applied Time Series Analysis for Managerial Forecasting*, San Francisco: Holden Day, 1973.

Trigg, D. W. "Monitoring a Forecasting System". *Operational Research Quarterly* 15(1964), pp. 271-274.

Trigg, D. W. y A. G. Leach. "Exponential Smoothing with Adaptive Response Rate". *Operational Research Quarterly* 18(1967), pp. 53-59.

Wiener, N. *Extrapolation*, *Interpolation*, *and Smoothing of Stationary Time Series*. Cambridge. MA: MIT Press, 1949.

Wilson, J. H. y B. Keating, *Business Forecasting*. Homewood, IL: Richard D. Irwin, 1990.

Winters, P. R. "Forecasting Sales by Exponentially Weighted Moving Averages". *Management Science* 6(1960), pp. 324-342.

Yule, G. U. "Why Do We Sometimes Get Nonsense Correlations between Time Series? A Study of Sampling and the Nature of Time Series". *Journal of the Royal Statistical Society* 89(1926), pp. 1-64.

Capítulo **Tres**

Planeación agregada

Panorama general del capítulo

Objetivo

Desarrollar técnicas para agregar unidades de producción y determinar los niveles adecuados de producción y de fuerza de trabajo con base en la demanda pronosticada de unidades agregadas.

Puntos clave

1. *Unidades agregadas de producción.* Este capítulo también podría denominarse planeación macro de la producción, puesto que el propósito de agregar unidades es poder desarrollar un plan integral para toda la compañía o para un subconjunto de la compañía, como la línea de producto o una planta particular. En el caso de las compañías grandes que producen una amplia variedad de productos, o de compañías que prestan servicios más que ofrecer productos, la determinación de las unidades agregadas puede representar un desafío. El enfoque más directo consiste en expresar unidades agregadas en alguna medida genérica, como dólares de ventas, toneladas de acero o galones de pintura. Para un servicio como el que proporciona una compañía o un despacho de abogados, las horas facturadas pueden ser una forma razonable de expresar unidades agregadas.
2. *Aspectos de la planeación agregada.* A continuación presentamos las características más importantes de la planeación agregada:
   - *Suavizamiento.* Costos que surgen al cambiar la producción y niveles de fuerza de trabajo.
   - *Cuellos de botella.* Planeación con anticipación a los periodos de demanda pico.
   - *Horizonte de planeación.* Debemos elegir con mucho cuidado el número de periodos considerados. Si son muy cortos, no es posible anticipar cambios en la demanda. Si son demasiado largos, los pronósticos de la demanda son poco confiables.
   - *Tratamiento de la demanda.* Todos los modelos matemáticos en este capítulo suponen que se conoce la demanda, es decir, tienen cero errores de pronóstico.
3. *Costos en la planeación agregada*
   - *Costos de suavizamiento.* El costo de cambiar la producción y/o los niveles de fuerza de trabajo.
   - *Costos de mantener el inventario.* El costo de oportunidad por los dólares invertidos en inventario.
   - *Costos de faltantes.* Los costos asociados con los pedidos atrasados o demanda perdida.
   - *Costos de mano de obra.* Esto incluye los costos de mano de obra directa en tiempo normal, tiempo extra, costos de subcontratación y costos de tiempo inactivo.
4. *Solución de problemas de planeación agregada.* Gráficamente, es posible encontrar soluciones aproximadas a los problemas de planeación agregada, así como la programación

lineal arroja soluciones exactas. Al momento de resolver gráficamente problemas, el primer paso es dibujar una gráfica de la curva de demanda neta acumulada. Si la meta es desarrollar un plan de nivel (es decir, uno que tenga producción constante o niveles de fuerza de trabajo sobre el horizonte de planeación), entonces se iguala la curva de la demanda neta acumulativa lo más cercanamente posible a una línea recta. Si la meta es desarrollar un plan de inventario cero (es decir, uno que minimice los costos de mantener un inventario y costos de faltantes), entonces se hace el seguimiento de la curva de la demanda neta acumulativa lo más cercanamente posible cada periodo. Aunque la programación lineal proporciona soluciones óptimas de costo, el método no toma en cuenta políticas de administración de cuentas, como puede ser evitar la contratación y el despido tanto como sea posible. Para un problema con un horizonte de planeación de T periodos, la formulación de programación lineal requiere de $8T$ variables y $3T$ restricciones. Para los horizontes de planeación largos, esto puede resultar sumamente tedioso. Otra cuestión que debe enfrentarse es que la solución de un programa lineal no sea entera. Para manejar este problema, tendríamos que especificar que las variables del problema son enteras (lo que computacionalmente lo haría difícil de manejar) o desarrollar un procedimiento de redondeo más adecuado.

5. *La regla de decisión lineal.* El concepto de planeación agregada tiene sus raíces en el trabajo de Holt, Modigliani, Muth y Simon (1960), quienes desarrollaron un modelo para Pittsburgh Paints (supuestamente) para determinar sus niveles de producción y fuerza de trabajo. El modelo utilizó aproximaciones cuadráticas para los costos y obtuvo ecuaciones lineales simples para las políticas óptimas. Este trabajo fue el origen del posterior interés en la planeación agregada.
6. *Modelado de la conducta administrativa.* Bowman (1963) consideró reglas de decisión lineal similares a aquellas deducidas por Holt, Modigliani, Muth y Simon, pero además sugirió ajustar los parámetros del modelo con base en las acciones de la administración, más que recomendar acciones óptimas basadas en la minimización de costos. Este es uno de los pocos ejemplos de modelo matemático usado para describir la conducta humana en el contexto de la planeación de operaciones.
7. *Desagregar planes agregados.* Aunque la planeación agregada sirve para proporcionar soluciones de planeación macro a nivel de la compañía, la pregunta es si estos planes agregados proporcionan alguna guía para la planeación en niveles inferiores. Un esquema de desagregación es un medio para tomar un plan agregado y descomponerlo en sus diferentes componentes para obtener planes más detallados en niveles más bajos dentro de la compañía.

A lo largo de la vida tomamos decisiones tanto micro como macro. Las decisiones micro pueden ser qué desayunaremos, qué ruta tomaremos para ir al trabajo, a cuál supermercado acudiremos o qué película rentaremos. Entre las decisiones macro se encuentran aquellas que pueden cambiar el curso de nuestra vida: dónde vivir, qué estudiar, dónde trabajar, con quién casarnos. Una compañía también debe tomar decisiones macro y micro todos los días: en este capítulo exploraremos las decisiones tomadas a nivel macro, como puede ser la planeación de los niveles de producción y de fuerza de trabajo en toda la compañía.

La **planeación agregada**, que también podría llamarse planeación macro de la producción, aborda el problema de decidir cuántos empleados debe ocupar una empresa; y, para una compañía manufacturera, la cantidad y mezcla de productos que debe producir. La planeación macro no se limita a las compañías manufactureras. Las organizaciones de servicios también deben determinar necesidades de personal. Por ejemplo, las aerolíneas deben planear los niveles de personal para sobrecargos y pilotos, y los hospitales los de enfermeras. Las estrategias de planeación macro son una parte fundamental en la estrategia general de negocios de la empresa. Algunas empresas trabajan con la filosofía de que los costos sólo se pueden controlar con cambios frecuentes de tamaño y/o composición de la fuerza de trabajo. La industria aeroespacial en California adoptó esa estrategia en la década de los setenta. A medida que los contratos

gubernamentales se desplazaron de un productor a otro, lo mismo ocurrió con la fuerza técnica de trabajo. Hay otras empresas que tienen fama de retener a sus empleados aun cuando los tiempos sean malos. Dos ejemplos recientes y bien conocidos eran IBM y AT&T.

La planeación macro comienza con el pronóstico de la demanda, ya sea que la empresa proporcione un servicio o fabrique un producto. Las técnicas para pronosticar la demanda se describieron en el capítulo 2. La velocidad de respuesta de la empresa respecto a los cambios previstos en la demanda depende de ciertos factores. Entre ellos está la estrategia general que pudiera tener en cuanto a conservar a sus empleados, y sus compromisos con los trabajadores actuales.

En el capítulo 2 vimos que los pronósticos de la demanda por lo general se equivocan porque casi siempre hay un componente aleatorio de ésta que no puede predecirse con exactitud. La metodología de planeación agregada que se describe en este capítulo reinicia con el supuesto de que la demanda es determinista, o que se conoce con anterioridad. Este supuesto se hace para simplificar el análisis y permitir concentrarnos en los cambios sistemáticos o predecibles, en el patrón de la demanda, y no en los cambios no sistemáticos, o aleatorios. En el capítulo 5 se describe con detalle la administración de inventarios sujeta a aleatoriedad.

Tradicionalmente, la mayoría de los manufactureros han optado por conservar la producción primaria para sí mismos. Pueden comprar algunos componentes a proveedores externos (véase la descripción del problema de fabricar o comprar en el capítulo 1), pero el producto primario por lo regular lo fabrica la empresa. Henry Ford fue uno de los primeros manufactureros estadounidenses en diseñar un negocio completamente integrado en dirección vertical. Incluso era propietario de una plantación de árboles de hule para no tener que comprar el caucho de sus neumáticos.

No obstante, ese modo de pensar está cambiando mucho. En los ambientes dinámicos, las empresas han visto que pueden ser más flexibles si la manufactura se consigue en el exterior, esto es, se subcontrata la mano de obra. Un ejemplo es Sun Microsystems, un productor californiano de estaciones de trabajo para cómputo. Sun, líder en el mercado, adoptó la estrategia de enfocarse en la innovación y diseño de sus productos más que en fabricarlos. Para ello ha establecido estrechas ligas al contratar fabricantes como Solectron Corporation (de San José, California), ganador del premio Baldrige de Calidad. Al subcontratar su función manufacturera primaria, Sun ha podido ser más flexible y enfocarse hacia la innovación en un mercado que cambia rápidamente.

La planeación agregada implica objetivos encontrados, como reaccionar con rapidez a cambios anticipados en la demanda, para lo que se requeriría hacer cambios frecuentes, y potencialmente grandes, en el tamaño de la fuerza de trabajo. A esa estrategia se le ha llamado estrategia de persecución. Puede ser rentable, pero también podría ser una mala estrategia empresarial a largo plazo. Los trabajadores que son despedidos quizás no estén disponibles cuando la empresa cambie. Por esta razón, tal vez la compañía prefiera adoptar el objetivo de retener una fuerza estable de trabajo. Sin embargo, así se incurre con frecuencia en grandes acumulaciones de inventario durante periodos de baja demanda. Las empresas de servicios pueden adquirir grandes deudas para cumplir con las nóminas en los periodos bajos. Un tercer objetivo es desarrollar un plan de producción para la empresa que maximice la ganancia en el horizonte de planeación sujeto a restricciones de capacidad. Los costos explícitos de efectuar los cambios deben ser factores clave en el proceso de decisión cuando el objetivo primario es la maximización de la ganancia.

El objetivo de la metodología de la planeación agregada es traducir los pronósticos de la demanda a un esquema de planeación para niveles de personal y de producción en la empresa, durante un horizonte predeterminado de planeación. Esa metodología no se limita a la planeación al máximo nivel. Aunque por lo general se considera que la planeación macro es una herramienta para determinar los niveles de fuerza de trabajo y de producción, las grandes empresas la pueden encontrar útil también a nivel de la planta. La planeación de la producción se puede considerar como un proceso jerárquico en el que las decisiones de compra, producción y personal se deben hacer en varios niveles de la empresa. Los métodos de planeación agregada pueden aplicarse casi en cualquier nivel, aunque el concepto consiste en administrar grupos de artículos, más que administrar artículos aislados. En los capítulos 4 y 5 describiremos los métodos de planeación de inventarios, cuya finalidad es controlar artículos aislados.

En este capítulo se analizan varias técnicas para determinar planes agregados. Algunos de ellos son heurísticos (aproximados) y otros óptimos. Esperamos transmitir al lector un concepto de los aspectos implicados en la planeación agregada, un conocimiento de las herramientas básicas disponibles para llegar a soluciones, y una idea de las dificultades asociadas al aplicar los planes agregados a situaciones del mundo real.

3.1 UNIDADES AGREGADAS DE PRODUCCIÓN

El método de planeación agregada está basado en la existencia de una unidad agregada de producción. Cuando la clase de artículos producidos es parecida, una unidad de producción agregada puede corresponder a un artículo "promedio", pero si se producen muchas clases distintas de artículos sería más adecuado considerar unidades agregadas en términos de peso (toneladas de acero), volumen (galones de gasolina), cantidad necesaria de trabajo (años-hombre de tiempo de programación), o valor monetario (valor del inventario en dólares). No siempre es obvio cuál esquema de planeación agregada debe ser el adecuado. Eso depende del contexto del problema particular de planeación y del nivel de agregación que se necesite.

Ejemplo *3.1*

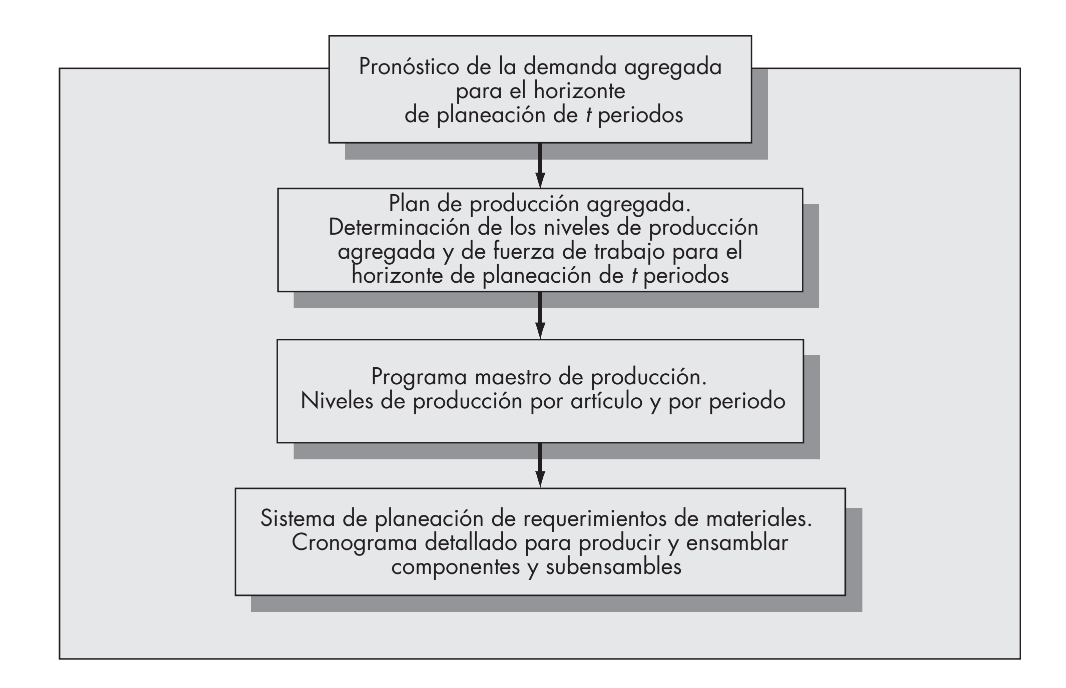
El gerente de planta de una gran empresa de electrodomésticos piensa implementar un sistema de planeación agregada para determinar los niveles de fuerza de trabajo y producción en su planta. Esta fábrica produce seis modelos de lavadoras. Sus características son:

Número de modelo	Cantidad de horas-hombre necesarias para producir	Precio de venta
A5532	4.2	$285
K4242	4.9	345
L9898	5.1	395
L3800	5.2	425
M2624	5.4	525
M3880	5.8	725

El gerente debe elegir un esquema particular de agregación. Una posibilidad es definir que la unidad agregada sea un dólar de producción. Desafortunadamente, los precios de venta de los diversos modelos de lavadora no son consistentes con la cantidad de horas-hombre requeridas para producirlos. La relación del precio de venta dividido entre las horas-hombre es 67.86 dólares para la A5532, y 125 dólares para la M3880. (La empresa basa sus precios en que los modelos menos costosos tienen mayor volumen de ventas.) El gerente observa que los porcentajes de la cantidad total de ventas para esos seis modelos han sido bastante constantes, dados los valores de 32% para la A5532, 21% para la K4242, 17% para la L9898, 14% para la L3800, 10% para la M2624 y 6% para la M3880. Entonces, decide definir una unidad agregada de producción como una lavadora ficticia que requiere $(.32)(4.2) + (.21)(4.9) + (.17)(5.1) + (.14)(5.2) + (.10)(5.4) + (.06)(5.8) = 4.856$ horas de mano de obra. El gerente puede obtener pronósticos de venta para unidades de producción agregada esencialmente de la misma manera, multiplicando las fracciones correspondientes por los pronósticos de unidades de venta de cada modelo de lavadora.

El método que utilizó el gerente de planta en el ejemplo 3.1 fue posible gracias a la relativa semejanza de los productos fabricados. Resulta más difícil, sin embargo, definir una unidad de producción agregada a un nivel superior de la empresa. En casos en los que se fabrica una gran variedad de productos, una unidad agregada natural es el dinero obtenido por ventas. Aunque, como vimos en el ejemplo, esto no se traducirá necesariamente en el mismo número de unidades de producción para cada artículo, por lo general proporciona una buena aproximación para planear el nivel máximo de una empresa que fabrica una línea diversificada de productos.

FIGURA 3-1
Jerarquía de las decisiones de planeación de la producción.



La planeación agregada (y el problema asociado de desagregar los planes agregados o convertirlos en programas maestros detallados) se vincula estrechamente con la planeación jerárquica de la producción (HPP, *hierarchical production planning*) enunciada por Hax y Meal (1975). En ella se consideran tamaños de fuerza de trabajo y tasas de producción en diversos niveles de la empresa, a diferencia de simplemente considerar el nivel máximo, como en la planeación agregada. Para fines de planeación agregada, Hax y Meal recomiendan la siguiente jerarquía:

1. *Artículos.* Son los productos finales que se entregarán al cliente. Con frecuencia a un artículo se le llama unidad de almacenamiento, UA, y representa el nivel de detalle más fino en la estructura del producto.
2. *Familias.* Se definen como un grupo de artículos que comparten un costo común de preparación de manufactura.
3. *Tipos.* Se trata de grupos de familias cuyas cantidades de producción se determinan con un solo plan agregado de producción.

En el ejemplo 3.1, los artículos corresponderían a los modelos individuales de lavadora. Una familia podría ser todas las lavadoras, y un tipo podría estar formado por todos los electrodomésticos grandes. El esquema de agregación de Hax-Meal no necesariamente funcionará en todos los casos. Por lo general, el modelo de agregación debe ser consistente con la estructura organizacional de la empresa y su línea de productos.

En la figura 3-1 se ilustra un esquema de la función de planeación agregada y su lugar en la jerarquía de las decisiones de planeación de la producción.

3.2 PANORAMA GENERAL DEL PROBLEMA DE PLANEACIÓN AGREGADA

Después de haber definido la unidad adecuada al nivel de la empresa para el que hay que determinar un plan agregado, suponemos que existe un pronóstico de la demanda para un horizonte específico de planeación, expresado en función de unidades de producción agregada. Sean $D_1, D_2, \ldots, D_T$ los pronósticos de demanda para los siguientes T periodos de planeación. En la mayoría de las aplicaciones, un periodo de planeación es un mes, aunque pueden desarrollarse planes agregados para otros intervalos de tiempo, como semanas, trimestres o años. Una

característica importante de la planeación agregada es que las demandas se consideran como constantes conocidas (es decir, se supone que el error del pronóstico es cero). Describiremos las razones para establecer este supuesto más adelante.

El objetivo de la planeación agregada es determinar las cantidades de producción agregada y los niveles requeridos de recursos para alcanzar esas metas de producción. En la práctica esto se traduce en precisar la cantidad de trabajadores que deben emplearse y la cantidad de unidades agregadas que se necesita producir en cada uno de los periodos de planeación 1, 2, . . . , T. El objetivo de la planeación agregada es balancear las ventajas de producir para cumplir con la demanda tan exactamente como sea posible y los problemas causados por el cambio de los niveles de producción y/o de la fuerza de trabajo.

Los principales aspectos relacionados con el problema de planeación agregada son:

1. *Suavizamiento.* Se refiere a los costos que resultan de cambiar los niveles de fuerza de trabajo de un periodo al siguiente. Dos de los componentes clave de los costos de suavizamiento son resultado de contratar y despedir trabajadores. La metodología de la planeación agregada requiere especificar esos costos, que pueden ser difíciles de estimar. Despedir a los trabajadores podría tener consecuencias y costos de gran alcance, que serían difíciles de evaluar. Las empresas que contratan y despiden con frecuencia desarrollan una mala imagen ante el público. Esto puede afectar las ventas con facilidad, desanimando a los empleados potenciales a unirse a la compañía. Además, los trabajadores que son despedidos podrían no estar dispuestos a esperar cerca de la compañía para conseguir un lugar. Despedir a los trabajadores puede tener un efecto negativo sobre el tamaño futuro de la fuerza laboral, si esos trabajadores obtienen empleo en otras industrias. Por último, la mayoría de las empresas simplemente no tienen libertad para contratar y despedir a voluntad. Las decisiones de la administración para alterar los niveles de fuerza de trabajo están restringidas por los convenios laborales.

2. *Problemas de cuello de botella.* El término *cuello de botella* se usa para indicar la incapacidad del sistema para responder a cambios repentinos en la demanda, que son resultado de restricciones en la capacidad. Por ejemplo, se puede producir un cuello de botella cuando el pronóstico de la demanda en un mes es extremadamente alto y la planta no tiene capacidad suficiente para cumplirla. Un cuello de botella puede ser ocasionado por una descompostura de un equipo vital, por ejemplo.

3. *Horizonte de planeación.* El número de periodos para los que se debe pronosticar la demanda, y por consiguiente la cantidad de periodos para los que se van a determinar los niveles de fuerza de trabajo y niveles de inventarios, debe especificarse por adelantado. La elección de T, el horizonte de pronóstico, puede ser importante para determinar la utilidad del plan agregado. Si T es demasiado pequeño, los niveles reales de producción podrían no ser adecuados para cumplir con la demanda más allá del horizonte. Si T es demasiado grande, es probable que los pronósticos en el futuro lejano sean imprecisos. Si las demandas futuras reales son muy diferentes de los pronósticos, las decisiones sugeridas por el plan agregado podrían ser incorrectas. Otro asunto donde interviene el horizonte de planeación es el efecto del *fin del horizonte.* Por ejemplo, el plan agregado podría recomendar que se redujera a cero el nivel de inventario al final del horizonte, para minimizar los costos de mantenerlo. Esta podría ser una mala estrategia, en especial si la demanda aumenta en ese momento. (No obstante, este problema en particular puede evitarse si se agrega una restricción que especifique niveles mínimos de inventario.)

En la práctica frecuentemente se utilizan programas progresivos. Esto significa que en el momento de la siguiente decisión se incorpora un nuevo pronóstico de la demanda a los anteriores, y que éstos deben modificarse para reflejar la información nueva. El nuevo plan agregado puede recomendar distintos niveles de producción y de fuerza de trabajo para el periodo actual, respecto a los sugeridos un periodo antes. El programa debe considerarse como dinámico más que estático cuando sólo se necesita implementar de inmediato las decisiones para el periodo actual de planeación.

Si bien los programas progresivos son comunes, es posible que debido a los tiempos de demora en la producción, el programa se deba congelar durante ciertos periodos de planeación. Esto quiere decir que las decisiones durante algunos periodos en el futuro no pueden alterarse. La forma más directa de manejar los horizontes congelados es simplemente designar como periodo 1 al primer periodo en el que las decisiones no estén congeladas.

4. *Tratamiento de la demanda.* Como ya se mencionó, en la metodología de la planeación agregada es necesario suponer que la demanda se conoce con certidumbre. Este es, a la vez, un punto débil y un punto fuerte de ese método. Es una debilidad porque no toma en cuenta la posibilidad (y de hecho, la probabilidad) de los errores de pronóstico. Como se dijo en la descripción de las técnicas de pronóstico en el capítulo 2, es virtualmente una certidumbre que los pronósticos de la demanda sean incorrectos. La planeación agregada no proporciona colchón alguno contra errores imprevistos en los pronósticos. Sin embargo, la mayoría de los modelos de inventarios que permiten demanda aleatoria requieren que la demanda promedio sea constante al paso del tiempo. La planeación agregada permite que el gerente se concentre en los cambios sistemáticos que por lo general no se asumen en modelos que suponen demanda aleatoria. Los efectos de las fluctuaciones estacionales y los ciclos económicos se pueden incorporar en la función de planeación cuando se supone una demanda determinista.

3.3 COSTOS EN PLANEACIÓN AGREGADA

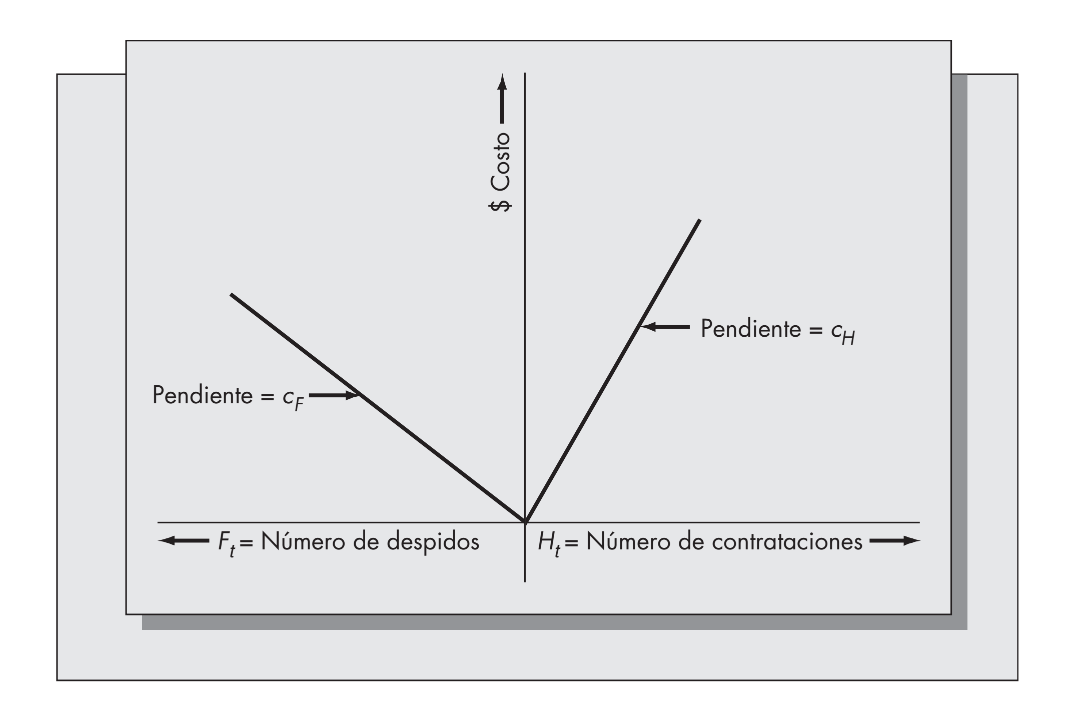
De la misma manera que en la mayoría de los problemas de optimización considerados por la administración de la producción, el objetivo del análisis es elegir el plan agregado que minimice costos. Es importante identificar y medir aquellos costos específicos que se vean afectados por la decisión de planeación.

1. *Costos de suavizamiento.* Son aquellos en los que se incurre al cambiar los niveles de producción de un periodo a otro. En el contexto de la planeación agregada, el costo más destacado de suavizamiento es el de cambiar el tamaño de la fuerza de trabajo. Su aumento requiere de tiempo y gastos para presentar los anuncios para los puestos, entrevistar a los probables empleados y capacitar a los contratados recientemente. Su disminución significa que hay que despedir trabajadores. Así, las indemnizaciones son el costo de disminuir el tamaño de la fuerza de trabajo. Otros costos, un poco más difíciles de medir, son: *a*) los costos de una declinación en la moral de los trabajadores, y *b*) el potencial de disminuir el tamaño de la oferta de trabajo en el futuro, como consecuencia de que los empleados despedidos se coloquen en otras empresas o industrias.

La mayoría de los modelos que se considerarán aquí suponen que los costos de aumentar y disminuir el tamaño de la fuerza de trabajo son funciones lineales de la cantidad de empleados que se contratan o despiden. Esto es, hay una cantidad monetaria constante que se asume por cada empleado que se contrata o despide. Es probable que el supuesto de linealidad sea razonable hasta cierto punto. Pero, a medida que escasea la oferta de trabajo, pueden requerirse costos adicionales para contratar más trabajadores, y los costos por despedir pueden incrementarse bastante si la cantidad de los despedidos es muy grande. En la figura 3-2 se ilustra una función característica del costo de cambiar el tamaño de la fuerza de trabajo.

2. *Costos por mantener inventarios.* Son aquellos en los que se incurre al tener capital invertido en los inventarios. Si la empresa logra disminuir su inventario, el dinero ahorrado podría ocuparse en otro rubro, con un rendimiento que varía según la industria y la empresa específica. (En el capítulo 4 se realiza una descripción más completa de los costos por mantener inventarios.) Casi siempre se supone que estos costos son lineales respecto a la cantidad de unidades que se mantienen en determinado punto en el tiempo. Para fines de la planeación agregada supondremos que el costo de mantener el inventario se expresa en unidades monetarias por unidad mantenida por periodo de planeación. También asumiremos que estos costos se cargan al inventario que permanece a mano al *final* del periodo de planeación. Este supuesto sólo se hace por comodidad, ya que podrían cargarse también al inventario inicial o al inventario promedio.

FIGURA 3-2
Costo de cambiar el tamaño de la fuerza de trabajo.



3. *Costos de faltantes.* Los costos por mantener inventarios se cargan contra el inventario agregado siempre que éste sea positivo. En algunos casos podrá ser necesario incurrir en faltantes, que se representan por un nivel negativo de inventario. Los faltantes se pueden presentar cuando la demanda pronosticada es mayor que la capacidad instalada para la producción, o cuando las demandas son mayores que las previstas. Para fines de planeación agregada, en general se supone que el exceso de demanda se programa y satisface en un periodo futuro. Sin embargo, en un ambiente muy competitivo, es posible que se pierda el exceso de demanda y que el cliente vaya a otro lado. Este caso, conocido como "ventas perdidas", es más adecuado en la administración de artículos únicos y se presenta con más frecuencia en un contexto de menudeo que en uno de manufactura.

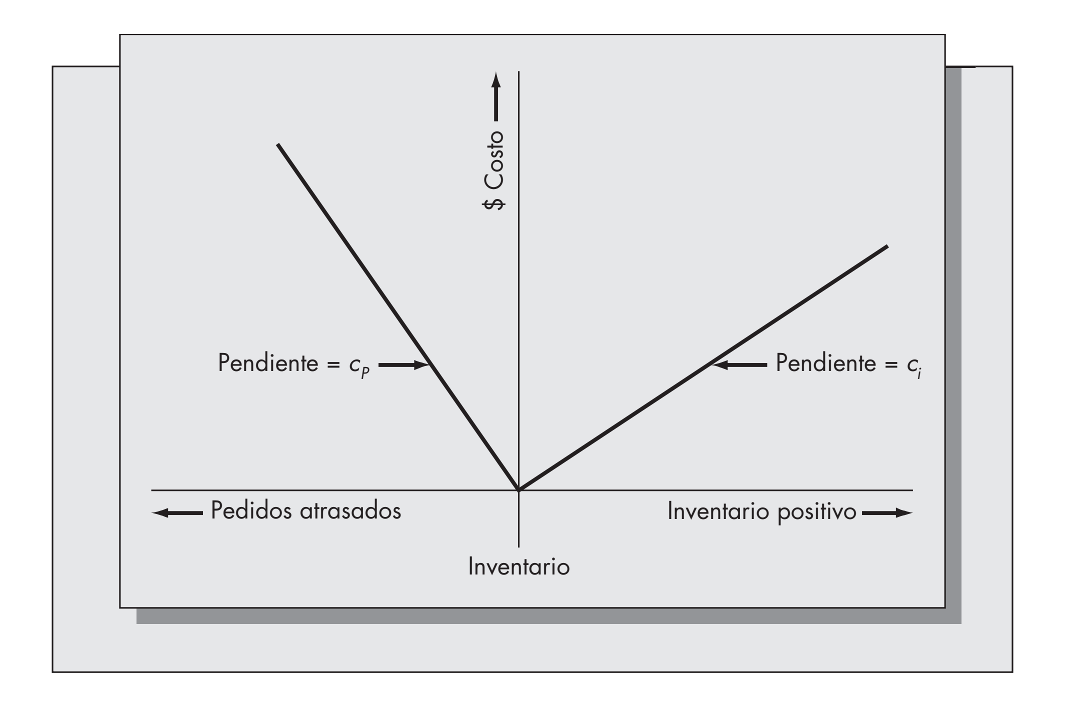
Generalmente, los costos por faltantes se suponen lineales (como ocurre con los costos por mantener inventarios). También se pueden describir con funciones convexas, pero las lineales parecen ser las más comunes. La figura 3-3 muestra una función característica del costo por mantener inventarios y el costo por faltantes de inventario.

4. *Costos de tiempo regular.* Son los costos por producir una unidad de producto durante horas normales de trabajo. En esta categoría se incluyen el costo real de la plantilla de empleados regulares que trabajan en tiempo normal, los costos directos e indirectos de materiales, así como otros gastos de manufactura. Cuando toda la producción se realiza en tiempo normal, los costos regulares de la nómina se transforman en "costo sumergido", porque la cantidad de unidades producidas debe ser igual a la de unidades demandadas durante cualquier horizonte de planeación de longitud suficiente. Los costos de la nómina usual no deben incluirse en la evaluación de las distintas estrategias si no hay tiempo extra ni trabajadores ociosos.

5. *Costos de tiempo extra* y *de subcontratación.* Son los costos de producción de las unidades que no se producen en el tiempo normal. El tiempo extra se refiere a producción con empleados de tiempo regular, pero más allá del horario normal de trabajo, y la subcontratación indica la producción de artículos por parte de un proveedor externo. De nuevo, se supone que por lo general esos costos son lineales.

6. *Costos de tiempo libre*. La formulación completa del problema de planeación agregada también debe incluir un costo por subutilización de la fuerza de trabajo o tiempo libre. En muchos contextos, el costo de tiempo libre es cero, ya que los costos directos de tiempo libre deben tomarse en cuenta en los costos de mano de obra y los niveles de producción más bajos. Sin embargo, el tiempo libre puede tener otras consecuencias para la compañía. Por ejemplo, si las

FIGURA 3-3
Costos por pedidos atrasados y por mantener inventarios.



unidades agregadas se introducen a otro proceso, el tiempo libre de la línea puede ocasionar costos más elevados para el proceso subsecuente. En tales casos, puede incluirse de manera explícita un costo de tiempo libre positivo.

Cuando la planeación se realiza en un nivel relativamente alto de la compañía, es más importante el efecto de los factores intangibles. Cualquier solución al problema de la planeación agregada obtenido de un modelo basado en costos debe considerarse con mucho cuidado en el contexto de las políticas de la compañía. Una solución óptima al modelo matemático puede provocar una política que requiera de una contratación y despido frecuentes de personal. Dicha política puede no ser factible, debido a contratos previos, o puede resultar indeseable, debido a los efectos negativos potenciales para la imagen pública de la empresa.

Problemas para las secciones 3.1 a 3.3

1. ¿Cuál es el significado de *unidad de producción agregada*? Las unidades de producción agregada, ¿corresponden siempre a artículos reales? ¿Alguna vez lo hacen? Explique su respuesta.
2. ¿Cuál es el esquema de agregación que recomiendan Hax y Meal?
3. Describa los términos siguientes y explique su relación con el problema de la planeación agregada.

   a) Suavizamiento.
   b) Cuellos de botella.
   c) Capacidad.
   d) Horizonte de planeación.

4. Un taller local emplea 60 trabajadores con diversas habilidades. El taller acepta pedidos únicos, y también mantiene varios clientes regulares. Describa algunas de las dificultades para el uso de la metodología de planeación agregada en este contexto.
5. Un gran fabricante de artículos para el hogar desea integrar un modelo de planeación agregada a su estrategia de manufactura. Dos de los vicepresidentes tienen opiniones totalmente contrarias acerca del valor del enfoque. ¿Qué argumentos usaría cada uno para respaldar su punto de vista?

6. Describa los siguientes costos y comente los problemas que surgen al tratar de medirlos en un ambiente operativo real.

   a) Costos de suavizamiento.
   b) Costos por mantener inventarios.
   c) Costos de nómina.

7. Comente la siguiente afirmación: "Desde que usamos un programa progresivo de producción, en realidad no necesitamos conocer la demanda más allá del mes siguiente".
8. El Saint Clair County Hospital trata de evaluar sus necesidades de enfermeras para los cuatro meses siguientes (de enero a abril). La necesidad de enfermeras depende de la cantidad y del tipo de pacientes en el hospital. A partir de un estudio que llevaron a cabo sus consultores, se ha determinado que se requieren las siguientes proporciones de enfermeras-pacientes:

Pacientes pronosticados	Cantidad de enfermeras requeridas por paciente	Tipo de paciente			
		Ene.	**Feb.**	**Mar.**	**Abr.**
Cirugía mayor	0.4	28	21	16	18
Cirugía menor	0.1	12	25	45	32
Maternidad	0.5	22	43	90	26
Cuidado intensivo	0.6	75	45	60	30
Otro	0.3	80	94	73	77

   a) ¿Cuántas enfermeras deben trabajar cada mes para coincidir lo mejor posible con los pronósticos de pacientes?
   b) Suponga que el hospital no quiere cambiar su política de no aumentar en más de 10% su cuerpo de enfermeras en cualquier mes. Recomiende un programa de plantilla de enfermeras durante los cuatro meses que cumpla tanto con ese requisito como con las necesidades de enfermeras cada mes.

3.4 UN PROBLEMA PROTOTIPO

De forma manual o con técnicas gráficas relativamente sencillas, pueden obtenerse soluciones adecuadas para muchos problemas de producción agregada. La programación lineal es un método para obtener soluciones (casi) óptimas. Las distintas técnicas de solución se ilustran por medio del siguiente ejemplo.

Ejemplo *3.2*

Densepack debe planear los niveles de fuerza laboral y de producción para un periodo de seis meses, de enero a junio. Esta empresa produce una línea de unidades de disco para computadoras compatibles con varias de las principales marcas. Las demandas pronosticadas durante los seis meses siguientes para determinada línea de unidades que se producen en la planta de Milpitas, California, son 1 280, 640, 900, 1 200, 2 000 y 1 400. Actualmente (al final de diciembre) hay 300 trabajadores en esa planta. Se espera que el inventario final en diciembre sea de 500 unidades, y la empresa desearía tener 600 unidades disponibles a finales de junio.

Hay varias formas de incorporar las restricciones de inventarios inicial y final en la formulación. La más cómoda consiste sólo en modificar los valores de la demanda pronosticada. La demanda neta pronosticada en el periodo 1 se define como la demanda predicha menos el inventario inicial. Si existe una restricción de inventario mínimo final, esa cantidad debe sumarse a la demanda en el periodo T. También se pueden manejar reservas mínimas de inventario si la demanda pronosticada se modifica. Si en

cada periodo hay un inventario mínimo de reserva, esa cantidad debe sumarse a la demanda en el primer periodo. Si en un periodo hay sólo un inventario mínimo de reserva, esa cantidad debe sumarse a la demanda de ese periodo y restar la del siguiente. No obstante, los inventarios finales reales deben calcularse usando el patrón original de la demanda.

Regresando a nuestro ejemplo, se definirá la demanda neta predicha para enero como igual a 780 (1 280 − 500), y la demanda neta predicha para junio como 2 000 (1 400 + 600). Al considerar la demanda neta se puede hacer un supuesto sencillo: que el inventario inicial y el final son iguales a cero. La demanda neta predicha y la demanda neta acumulada para los seis meses de enero a junio se muestran a continuación:

Mes	**Demanda neta pronosticada**	**Demanda neta acumulada**
Enero	780	780
Febrero	640	1 420
Marzo	900	2 320
Abril	1 200	3 520
Mayo	2 000	5 520
Junio	2 000	7 520

En la figura 3-4 se ilustra la demanda neta acumulada. Un plan de producción es la especificación de los niveles de producción para cada mes. Si no se permiten faltantes, la producción acumulada debe ser, como mínimo, igual a la demanda acumulada en cada periodo. Además de la demanda neta acumulada, la figura 3-4 también muestra un plan factible de producción.

Con el fin de ilustrar los intercambios con los costos de varios planes de producción supondremos en el ejemplo anterior que sólo hay tres costos por considerar: el de contratar trabajadores, el de despedir trabajadores y el de mantener el inventario. Definiremos

C_H = Costo de contratar un trabajador = \$500.

C_F = Costo de despedir un trabajador = \$1 000.

C_I = Costo de mantener una unidad de inventario durante un mes = \$80.

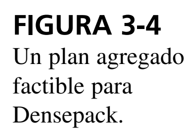
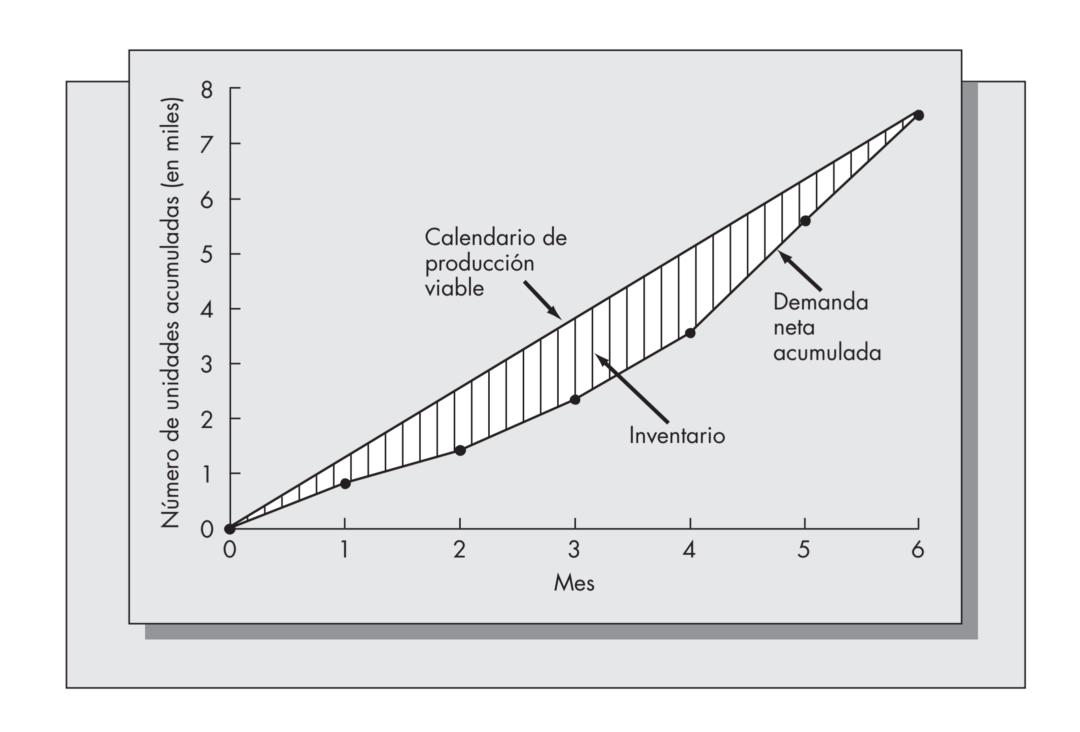


FIGURA 3-4
Un plan agregado factible para Densepack.



Se requiere un medio para traducir la producción agregada en unidades a niveles de fuerza de trabajo. Como no todos los meses tienen una cantidad igual de días hábiles, usaremos un día como unidad indivisible para medir, y definiremos

K = Número de unidades agregadas producidas por un trabajador en un día.

En el pasado, el gerente de planta observó que durante 22 días hábiles, con un nivel de fuerza de trabajo constante de 76 trabajadores, su fábrica produjo 245 unidades de disco. Esto quiere decir que, en promedio, la tasa de producción fue $245/22 = 11.1364$ unidades por día, cuando había 76 trabajadores trabajando en la planta. Por consiguiente, un trabajador produjo un promedio de $11.1364/76 = 0.14653$ unidades en un día. Así, $K = 0.14653$ para este ejemplo.

Analizaremos dos planes alternativos para administrar la fuerza de trabajo que representan dos estrategias administrativas esencialmente opuestas. El plan 1 es cambiar la fuerza de trabajo cada mes con el fin de producir las unidades suficientes para coincidir lo más posible con el patrón de la demanda. Esto se conoce como *plan de cero inventarios.* El plan 2 es mantener la fuerza de trabajo mínima para satisfacer la demanda neta. Ésta se conoce como *plan de fuerza de trabajo constante.*

Evaluación de una estrategia de persecución (plan de cero inventarios)

Ahora desarrollaremos un plan de producción para Densepack que minimice los niveles de inventarios que debe mantener la compañía durante el horizonte de planeación a seis meses. En la tabla 3-1 se resume la información necesaria para los cálculos, y se muestra la cantidad mínima de trabajadores requerida cada mes.

Para obtener las entradas de la columna final de la tabla 3-1, que son las cantidades mínimas de trabajadores requeridos cada mes, se divide la demanda neta pronosticada entre la cantidad de unidades producida por trabajador. El valor de esta relación se redondea *hacia arriba*, hasta el entero siguiente. Debe redondearse así para garantizar que no se presente ningún faltante. Por ejemplo, veamos el mes de enero. Al calcular la relación 780/2.931 se obtiene 266.12, que se redondea a 267 trabajadores. El número de días hábiles en cada mes depende de varios factores, como vacaciones pagadas y programas de mano de obra. La cantidad reducida de días en junio se debe a un paro programado de la planta la última semana de junio.

Recuérdese que la cantidad de trabajadores empleados al final de diciembre es 300. Al contratar y despedir trabajadores cada mes para coincidir con la demanda tanto como sea posible, se obtiene el plan agregado que muestra la tabla 3-2.

TABLA 3-1 **Cálculos iniciales para el plan de cero inventarios en Densepack**

A	B	C	D	E
Mes	Número de días de trabajo	Número de unidades producidas por trabajador (B × 0.14653)	Pronóstico de demanda neta	Número mínimo de trabajadores requeridos (D/C redondeado hacia arriba)
Enero	20	2.931	780	267
Febrero	24	3.517	640	182
Marzo	18	2.638	900	342
Abril	26	3.810	1 200	315
Mayo	22	3.224	2 000	621
Junio	15	2.198	2 000	910

TABLA 3-2 **Plan agregado de cero inventarios para Densepack**

A	B	C	D	E	F	G	H	I
Mes	Número de trabajadores	Número de contrataciones	Número de despidos	Número de unidades por trabajador	Número de unidades producidas (B × E)	Producción acumulada	Demanda acumulada	Inventario final (G − H)
Enero	267		33	2.931	783	783	780	3
Febrero	182		85	3.517	640	1 423	1 420	3
Marzo	342	160		2.638	902	2 325	2 320	5
Abril	315		27	3.810	1 200	3 525	3 520	5
Mayo	621	306		3.224	2 002	5 527	5 520	7
Junio	910	289		2.198	2 000	7 527	7 520	7
Totales		755	145					30

El número de unidades producidas cada mes (columna F en la tabla 3-2) se obtiene con la siguiente fórmula:

$$\begin{array}{l}\text{Número de}\\\text{unidades producidas}\end{array} = \begin{array}{l}\text{Número de}\\\text{trabajadores}\end{array} \times \begin{array}{l}\text{Número promedio de unidades agregadas}\\\text{que produce un solo trabajador en un mes}\end{array}$$

redondeado al entero superior siguiente.

Al multiplicar los totales de la parte inferior de la tabla 3-2 por los costos correspondientes se obtiene el costo total de este plan de producción. Para este ejemplo, el costo total de contratar, despedir y mantener es (755)(500) + (145)(1 000) + (30)(80) = $524 900. Ahora este costo debe ajustarse para incluir el costo por mantener el inventario final de 600 unidades, que fue excluido de la demanda para junio. Por consiguiente, el costo total de este plan es 524 900 + (600)(80) = $572 900. Obsérvese que el inventario inicial de 500 unidades no entra en los cálculos, porque desaparecerá durante el mes de enero.

Generalmente es imposible lograr cero inventarios al final de cada periodo de planeación, puesto que no se puede emplear una cantidad fraccionaria de trabajadores. Por esta razón, casi siempre habrá inventario remanente al final de cada periodo, además del que debe estar disponible al final del horizonte de planeación.

Es posible que el inventario final se pueda acumular en uno o más periodos hasta un punto en que se pueda reducir la fuerza de trabajo en uno o más trabajadores. En este ejemplo hay suficientes existencias disponibles para reducir la fuerza de trabajo en un trabajador en los meses de marzo y mayo. Compruebe que con el plan resultante se contrata un total de 753 trabajadores y se despide un total de 144 trabajadores, y sólo se tiene un total de 13 unidades de inventario. Este plan modificado tiene un costo de $569 540.

Evaluación del plan de fuerza de trabajo constante

El objetivo consiste ahora en eliminar por completo la necesidad de contratar y despedir durante el horizonte de planeación. Para garantizar que no se presenten faltantes de inventario en cualquier periodo es necesario calcular la fuerza mínima de trabajo requerida para *cada* mes en el horizonte de planeación. La demanda neta acumulada para enero es 780, y hay 2.931 unidades producidas por trabajador, lo que da como resultado una fuerza de trabajo mínima igual a 267 en enero. Hay exactamente 2.931 + 3.517 = 6.448 unidades producidas por trabajador en enero y febrero combinados, para los cuales la demanda acumulada es 1 420. Por consiguiente, se requieren 1 420/6.448 = 220.22 ≈ 221 trabajadores para cubrir enero y febrero. Continuando con las relaciones de la demanda neta acumulada y la cantidad acumulada de unidades producidas por trabajador para cada mes en el horizonte, se obtiene la tabla 3-3.

TABLA 3-3 **Cálculo de la fuerza de trabajo mínima requerida por Densepack**

A	B	C	D
Mes	**Demanda neta acumulada**	**Número acumulado de unidades por trabajador**	**Relación B/C (redondeada hacia arriba)**
Enero	780	2.931	267
Febrero	1 420	6.448	221
Marzo	2 320	9.086	256
Abril	3 520	12.896	273
Mayo	5 520	16.120	343
Junio	7 520	18.318	411

TABLA 3-4 **Niveles de inventario para el programa de fuerza de trabajo constante**

A	B	C	D	E	F
Mes	**Número de unidades producidas por trabajador**	**Producción mensual (B × 411)**	**Producción acumulada**	**Demanda neta acumulada**	**Inventario final (D − E)**
Enero	2.931	1 205	1 205	780	425
Febrero	3.517	1 445	2 650	1 420	1 230
Marzo	2.638	1 084	3 734	2 320	1 414
Abril	3.810	1 566	5 300	3 520	1 780
Mayo	3.224	1 325	6 625	5 520	1 105
Junio	2.198	903	7 528	7 520	8
Total					5 962

El número mínimo de trabajadores requeridos para el periodo completo de planeación es el elemento máximo en la columna D, en la tabla 3-3, que es 411 trabajadores. Haber obtenido la relación máxima en el periodo final sólo fue una coincidencia.

Ya que hay 300 trabajadores empleados al final de diciembre, para el plan de fuerza de trabajo constante se requiere contratar 111 trabajadores al principio de enero. No es necesario hacer más contrataciones ni despidos. Los niveles de inventario que resultan con una fuerza de trabajo constante igual a 411 trabajadores son los que aparecen en la tabla 3-4. Los niveles mensuales de producción, en la columna C de esa tabla, se obtienen multiplicando la cantidad de unidades producidas por trabajador cada mes, por el tamaño fijo de fuerza de trabajo, 411 trabajadores. El total de los niveles finales de inventario es 5 962 + 600 = 6 562. (Recuerde que se eliminaron las 600 unidades de la demanda para junio.) Por consiguiente, el costo total de inventario para este plan es (6 562)(80) = $524 960. A esto se agrega el costo de aumentar la plantilla de personal de 300 a 411 en enero, que es (111)(500) = $55 500, lo que da un costo total de $580 460 para este plan, ligeramente mayor que el costo del plan con cero inventarios, que fue $569 540. Sin embargo, como los costos de los dos planes son cercanos entre sí, es probable que, para eliminar cualquier costo no previsto por hacer cambios frecuentes en su fuerza de trabajo, la empresa prefiera el de la fuerza de trabajo constante.

Estrategias mixtas y restricciones adicionales

Las estrategias de los planes de cero inventarios y de la fuerza de trabajo constante que hemos descrito son estrategias puras: tienen el fin de lograr un objetivo. Con más flexibilidad, al introducir pequeñas modificaciones se pueden obtener costos bastante menores. El lector podría

preguntarse si tienen interés los cálculos manuales, considerando que los problemas de planeación agregada pueden formularse y resolverse en forma óptima por medio de la programación lineal.

A este respecto, los cálculos manuales estimulan la intuición y la comprensión. Es fácil pasar por alto una restricción u objetivo críticos cuando se usa la computadora. Es importante tener un sentido de la solución correcta antes de resolver un problema con ayuda de la computadora, para que los errores grandes sean obvios. Una habilidad, casi olvidada en nuestra época, es formular cálculos mentalmente antes de echar mano de la calculadora o la computadora.

La estrategia de fuerza de trabajo constante de Densepack se muestra en la figura 3-4. El área delimitada representa el inventario en existencia cada mes. Suponga que permitimos un solo cambio en la tasa de producción durante los seis meses. A partir de la tabla, ¿puede usted identificar una estrategia que reduzca el inventario en forma sustancial sin permitir faltantes?

Gráficamente, el problema radica en cubrir la curva de la demanda neta acumulada con dos rectas, no con una sola. Esto puede lograrse forzando que el inventario neto desaparezca al final del periodo 4 (abril). Se necesita suficiente producción en cada uno de los meses de enero hasta abril para satisfacer la demanda neta acumulada cada mes. Esto quiere decir que debemos producir $3\,520/4 = 880$ unidades en cada uno de los primeros cuatro meses. En la figura 3-4, la recta que une el origen con la demanda neta acumulada en abril queda totalmente arriba de la curva de demanda neta acumulada para los primeros meses. Si la gráfica es exacta, eso significa que no debe haber faltantes durante esos meses. Después, durante mayo y junio la producción se establece en 2 000 y coincide exactamente con la demanda neta respectiva. Con esa política se obtiene lo siguiente:

Mes	Demanda neta acumulada	Producción acumulada
Enero	780	880
Febrero	1 420	1 760
Marzo	2 320	2 640
Abril	3 520	3 520
Mayo	5 520	5 520
Junio	7 520	7 520

En la sección 3.5 veremos que esta política será la óptima para el problema de Densepack.

Puede usarse también el método gráfico de solución cuando hay más restricciones presentes. Por ejemplo, suponga que la capacidad de producción de la planta sólo es de 1 800 unidades por mes. Entonces, la anterior política no es factible en mayo y junio. En este caso, la restricción significa que la pendiente de la curva de producción acumulada está acotada en 1 800. Una solución sería producir 980 en cada uno de los primeros cuatro meses, y 1 800 unidades en cada uno de los dos últimos meses. Otra restricción podría ser que el cambio máximo de un mes al siguiente no superara las 750 unidades. Recomiende un plan de producción que cumpla con esta restricción.

Los anteriores son algunos ejemplos de restricciones que pueden surgir al usar la metodología de planeación agregada. A medida que las restricciones se hacen más complicadas, la determinación gráfica de buenas soluciones se dificulta cada vez más. Afortunadamente, la mayoría de las restricciones de este tipo pueden incorporarse facilmente en las formulaciones de programación lineal de los problemas de planeación agregada.

Problemas para la sección 3.4

9. Harold Grey posee una huerta pequeña en el Salinas Valley, donde cosecha albaricoques. Dichas frutas se secan en la huerta y se venden a varias cadenas de supermercados. Con base en la experiencia y en los contratos comprometidos, calcula que durante los cinco años siguientes las ventas, en miles de paquetes, serán:

Año	Demanda pronosticada (miles de paquetes)
1	300
2	120
3	200
4	110
5	135

Suponga que cada empleado permanece en su puesto durante al menos un año, y que en la actualidad Grey tiene tres trabajadores en su nómina. Grey estima que tendrá 20 000 paquetes disponibles al final del presente año. Considere que a cada trabajador se le pagan $25 000 en promedio al año, y que la responsabilidad de cada uno es producir 30 000 paquetes. Los costos de inventario son 4 centavos anuales por paquete, y no se permiten faltantes del producto.

Con base en su experiencia entrevistando y entrenando a nuevos trabajadores, Grey calcula que cada trabajador contratado cuesta $500. La indemnización por trabajador despedido es de $1 000.

a) Suponga que no se permiten faltantes de producto en inventario y calcule la fuerza de trabajo constante mínima que necesitará durante los cinco años siguientes.

b) Evalúe el costo del plan propuesto en el inciso *a*).

10. Grafique la demanda neta acumulada para los datos del problema 9.

a) Determine por medio de una gráfica un plan de producción que cambie la tasa de producción exactamente una vez durante los cinco años, y evalúe el costo de ese plan.

b) Determine por medio de una gráfica un plan de producción que cambie la tasa de producción exactamente dos veces, durante los cinco años, y evalúe el costo de ese plan.

11. Un supuesto implícito realizado en el problema 9 fue que los albaricoques secos no vendidos al final del año se pueden vender en los años siguientes. Digamos que los albaricoques no vendidos al final de cualquier año deben desecharse. Supongamos que el costo de desecharlos es $0.20 por paquete. Resuelva el problema 9 bajo estas condiciones.

12. El departamento de personal de A&M Corporation desea saber cuántos trabajadores se necesitarán cada mes durante el siguiente periodo de producción de seis meses. A continuación vemos un pronóstico mensual de la demanda para ese periodo.

Mes	Demanda pronosticada
Julio	1 250
Agosto	1 100
Septiembre	950
Octubre	900
Noviembre	1 000
Diciembre	1 150

A finales de junio el inventario disponible fue 500 unidades. La compañía desea mantener un inventario mínimo de 300 unidades cada mes, y quisiera tener 400 unidades disponibles al final de diciembre. Para producir cada unidad se requieren cinco horas-empleado, hay 20 días de trabajo cada mes, y cada empleado trabaja ocho horas cada día. La fuerza de trabajo al final de junio era de 35 trabajadores.

a) Determine el plan de producción con inventarios mínimos (es decir, uno que permita contratación y despidos arbitrarios).

b) Determine el plan de producción que cumpla con la demanda, pero que no implique la contratación ni el despido de trabajadores durante ese periodo de seis meses.

13. Mr. Meadows Cookie Company elabora diversas galletas de chocolate en su planta de Albion, Michigan. A partir de los pedidos recibidos y de los pronósticos de hábitos de compra, se estima que la demanda para los cuatro meses siguientes será 850, 1 260, 510 y 980, en miles de galletas. Durante un periodo de 46 días, cuando había 120 trabajadores, la empresa produjo 1.7 millones de galletas. Suponga que la cantidad de días hábiles en los cuatro meses son, respectivamente, 26, 24, 20 y 16. Actualmente hay 100 trabajadores empleados y no hay inventario inicial de galletas.

a) ¿Cuál es la fuerza de trabajo mínima constante requerida para cumplir con la demanda de los siguientes cuatro meses?

b) Suponga que c_I = 10 centavos por galleta por mes, c_H = \$100, y c_F = \$200. Evalúe el costo del plan propuesto en el inciso *a*).

14. Superchip, una empresa local de semiconductores, planea sus niveles de personal y producción para el año próximo. La compañía fabrica diversos microprocesadores, y usa las ventas en dólares para medir su producción agregada. Con base en los pedidos recibidos y en los pronósticos de venta proporcionados por el departamento de mercadotecnia, se estima que el año siguiente las ventas mensuales en dólares serán:

Mes	Días hábiles	Demanda pronosticada (en $10 000)
Enero	22	340
Febrero	16	380
Marzo	21	220
Abril	19	100
Mayo	23	490
Junio	20	625
Julio	24	375
Agosto	12	310
Septiembre	19	175
Octubre	22	145
Noviembre	20	120
Diciembre	16	165

Los costos por mantener inventarios se basan en un cargo anual de 25% de interés. Se anticipa que habrá 675 trabajadores en la nómina al final del año actual, y que los inventarios equivaldrán a $120 000. La empresa quisiera tener cuando menos $100 000 de inventario al final de diciembre del año próximo. Se estima que cada trabajador produce artículos que equivalen a $60 000 al año (suponga que un año consiste en 250 días hábiles). El costo de contratar un trabajador nuevo es de $200, y de despedir uno $400.

a) Determine la fuerza de trabajo constante mínima que cumpliría con la demanda predicha para el año siguiente.
b) Evalúe el costo del plan determinado en el inciso *a*).

15. Para los datos del problema 14, determine el costo del plan que cambia el tamaño de la fuerza de trabajo en cada periodo para cumplir lo más exactamente posible con la demanda.
16. Grafique la demanda neta acumulada para los datos de Superchip del problema 14. Por medio de una gráfica determine un plan de producción que no cambie los niveles de producción más de tres veces, y calcule el costo total de ese plan.

3.5 SOLUCIÓN DE PROBLEMAS DE PLANEACIÓN AGREGADA MEDIANTE PROGRAMACIÓN LINEAL

Programación lineal es un concepto que se utiliza para describir una clase general de problemas de optimización. El objetivo es determinar valores de n variables reales no negativas para maximizar o minimizar una función lineal de dichas variables, sujeta a m restricciones lineales, de las mismas variables.[1] La ventaja principal de formular un problema como programa lineal es que se pueden encontrar las soluciones óptimas eficazmente con el método simplex.

Cuando todas las funciones de costo son lineales, existe una formulación del problema general de planeación agregada como programación lineal. Se pueden obtener soluciones (esencialmente) óptimas para problemas muy grandes, gracias a la eficiencia de los programas de cómputo comerciales.[2]

Parámetros de costo e información dada

Asumiendo que los siguientes valores se conocen:

c_H = Costo de contratar un trabajador.
c_F = Costo de despedir un trabajador.
c_I = Costo de mantener una unidad de inventario durante un periodo.
c_R = Costo de producir una unidad en tiempo normal.
c_O = Costo incremental de producir una unidad en tiempo extra.
c_U = Costo de inactividad, por unidad de producción.
c_S = Costo de subcontratar una unidad de producción.
n_t = Número de días de producción en el periodo t.
K = Número de unidades agregadas producidas por un trabajador en un día.
I_0 = Inventario en existencia al inicio del horizonte de planeación.
W_0 = Fuerza de trabajo al inicio del horizonte de planeación.
D_t = Pronóstico de la demanda en el periodo t.

Los parámetros de costo también pueden ser funciones del tiempo, esto es, pueden cambiar con t. De esta forma podrían ser útiles para modelar cambios en los costos de contratación o despido debidos, por ejemplo, a una escasez en la oferta de mano de obra, o cambios en los costos de producción y/o almacenamiento debidos a insuficiencia en el suministro de recursos, o a cambios en las tasas de interés.

[1] Se puede encontrar un panorama general de la programación lineal en el suplemento después de este capítulo.

[2] Se incluye el calificativo porque con el redondeo se pueden obtener soluciones subóptimas. Más adelante regresaremos a este punto.

Variables del problema

Las variables del problema son:

W_t = Nivel de la fuerza de trabajo en el periodo t.
P_t = Nivel de producción en el periodo t.
I_t = Nivel de inventario en el periodo t.
H_t = Cantidad de trabajadores contratados en el periodo t.
F_t = Cantidad de trabajadores despedidos en el periodo t.
O_t = Producción en unidades durante el tiempo extra.
U_t = Tiempo inactivo de los trabajadores, en unidades de producción.
S_t = Cantidad de unidades subcontratadas en el exterior.

Las variables de tiempo extra y tiempo inactivo se determinan de la siguiente manera. El término Kn_t representa la cantidad de unidades producidas por un trabajador en el periodo t, y así Kn_tW_t sería la cantidad de unidades producidas por toda la fuerza de trabajo en el periodo t. Sin embargo, no es necesario que $Kn_tW_t = Pt$. Si $Pt > Kn_tW_t$, la cantidad de unidades producidas es mayor que lo que la fuerza de trabajo puede producir en tiempo normal. Esto significa que la diferencia se produce en tiempo extra, y entonces la cantidad de unidades producidas en tiempo extra es, exactamente, $O_t = P_t - Kn_tW_t$. Si $P_t < Kn_tW_t$, la fuerza de trabajo produce menos de lo que debería en tiempo normal, y eso significa que hay tiempo inactivo de trabajadores. El tiempo inactivo se mide en unidades de producción, no en tiempo, y está dado por $U_t = Kn_tW_t - P_t$.

Restricciones del problema

En la formulación de programación lineal se necesitan tres conjuntos de restricciones. Se incluyen para asegurar que se cubran la conservación de la fuerza de trabajo y la de unidades.

1. Restricciones de conservación de la fuerza de trabajo.

W_t	=	W_{t-1}	+	H_t	−	F_t	para $1 \le t \le T$.
Número de trabajadores en t	=	Número de trabajadores en $t - 1$	+	Número de despidos en t	−	Número de despidos en t	

2. Restricciones de conservación de unidades

I_t	=	I_{t-1}	+	P_t	+	S_t	−	D_t	para $1 \le t \le T$.
Inventario en t	=	Inventario en $t - 1$	+	Número de unidades producidas en t	+	Número de unidades subcontratadas en t	−	Demanda en t	

3. Restricciones que relacionan los niveles de producción con los niveles de fuerza.

P_t	=	Kn_tW_t	+	O_t	−	U_t	para $1 \le t \le T$.
Número de unidades producidas en t	=	Número de unidades producidas por fuerza de trabajo regular en t	+	Número de unidades producidas en tiempo extra en t	−	Número de unidades de producción inactiva en t	

Además de estas restricciones, la programación lineal requiere que todas las variables del problema sean no negativas. Estas restricciones y las de no negatividad son lo mínimo que debe estar presente en cualquier formulación. Observe que los puntos (1), (2) y (3) equivalen a $3T$ restricciones, y no a tres restricciones, donde T es la longitud del horizonte de pronóstico.

La formulación también requiere especificar el inventario inicial, I_0, y la fuerza de trabajo inicial, W_0, y puede incluir la especificación del último inventario en el periodo final, I_T.

Todos los costos antes definidos están incluidos en la función objetivo. La formulación de programación lineal consiste en elegir los valores de las variables en el problema, que en este caso son W_t, P_t, I_t, H_t, R_t, O_t, U_t y S_t para

$$\text{Minimizar} \sum_{t=1}^{T} (c_H H_t + c_F F_t + c_I I_t + c_R P_t + c_O O_t + c_U U_t + c_S S_t)$$

sujeta a

$$W_t = W_{t-1} + H_t - F_t \qquad \text{para } 1 \le t \le T \quad \text{(conservación de la fuerza de trabajo),} \tag{A}$$

$$P_t = Kn_t W_t + O_t - U_t \qquad \text{para } 1 \le t \le T \quad \text{(producción y fuerza de trabajo)} \tag{B}$$

$$I_t = I_{t-1} + P_t + S_t - D_t \qquad \text{para } 1 \le t \le T \text{ (balance de inventario),} \tag{C}$$

$$H_t, F_t, I_t, O_t, U_t, S_t, W_t, P_t \ge 0 \ \text{ (no negatividad),} \tag{D}$$

más cualesquiera restricciones adicionales que definan los valores del inventario inicial, fuerza de trabajo inicial, inventario final u otras variables cuyos valores hayan sido fijados por adelantado.

Redondeo de las variables

Los valores óptimos de las variables del problema, por lo general, no serán enteros. Sin embargo, los valores fraccionarios de muchas de las variables no tienen sentido. Se incluyen entre éstas el tamaño de la fuerza de trabajo, la cantidad de trabajadores contratados y la de trabajadores despedidos cada periodo, y también puede incluir la cantidad de unidades producidas cada periodo. (Es posible que en algunas aplicaciones se produzcan cantidades fraccionarias de unidades.) Una forma de manejar este problema es pedir por adelantado que algunas o todas las variables del problema sólo asuman valores enteros. Desafortunadamente, esto complica bastante el algoritmo de solución. Se requiere un mayor esfuerzo computacional para solucionar el problema resultante, conocido como problema de programación lineal entera, si lo comparamos con la programación lineal ordinaria. Es claro que resolver los problemas de tamaño moderado con programación lineal entera es una alternativa razonable.

Cuando no se cuenta con un software de programación entera, o cuando simplemente el problema es demasiado grande como para resolverlo de esa forma, la programación lineal ofrece una solución manejable. No obstante, después de obtenerla, algunas de las variables del problema deben redondearse a valores enteros. Un simple redondeo de cada variable hasta el entero más próximo puede conducir a una solución no factible y/o a una en la que sean inconsistentes los niveles de producción y la fuerza de trabajo. Recomendamos el siguiente método conservador: redondear los valores de las cantidades de trabajadores, en cada periodo t, a W_t, el entero siguiente mayor. Una vez determinados los valores de W_t se pueden calcular los valores de H_t, F_t y P_t junto con el costo del plan resultante.

El redondeo conservador siempre dará como resultado una solución posible, o factible, pero casi nunca llegará a la solución óptima. La solución conservadora se puede mejorar, por lo general, con cálculos de tanteo.

Si se logra formular un problema como programa lineal nada garantiza que la solución final tenga sentido en el contexto del problema. En el problema de planeación agregada no tiene sentido que haya producción en tiempo extra a la vez que exista tiempo inactivo en el mismo periodo, ni tampoco que se deba contratar y despedir trabajadores en el mismo lapso. Esto significa que O_t y U_t, una de esas variables, o ambas, deben ser cero, y que una o ambas H_t y F_t deben ser cero para cada t, $1 \leq t \leq T$. Este requisito se puede incluir explícitamente al formular el problema agregando las restricciones

$$O_tU_t = 0 \qquad \text{para } 1 \leq t \leq T,$$
$$H_tF_t = 0 \qquad \text{para } 1 \leq t \leq T,$$

porque si el producto de dos variables es cero, quiere decir que al menos una de ellas debe ser igual a cero. Desafortunadamente, esas restricciones no son lineales, porque implican un producto de variables del problema. Sin embargo, se ha visto que no es necesario incluir esas restricciones en forma explícita, porque la solución óptima de un problema de programación lineal siempre se da en un punto extremo de la región posible. Se puede demostrar que toda solución de producto extremo tiene esta propiedad de manera automática. Si no fuera así, la solución por programación lineal no tendría sentido.

Extensiones

También se puede usar la programación lineal para resolver versiones algo más generales del problema de planeación agregada. Se puede tener en cuenta la incertidumbre de la demanda en forma indirecta, suponiendo que hay un inventario mínimo de reserva B_t cada periodo. En ese caso se incluirían las restricciones

$$I_t \geq B_t \qquad \text{para } 1 \leq t \leq T.$$

Las constantes B_t tendrían que especificarse por adelantado. En forma semejante se podrían incluir cotas superiores a la cantidad de trabajadores contratados y a la de trabajadores despedidos cada periodo. Las restricciones sobre la cantidad de producción en cada periodo pueden representarse con facilidad mediante el conjunto de restricciones:

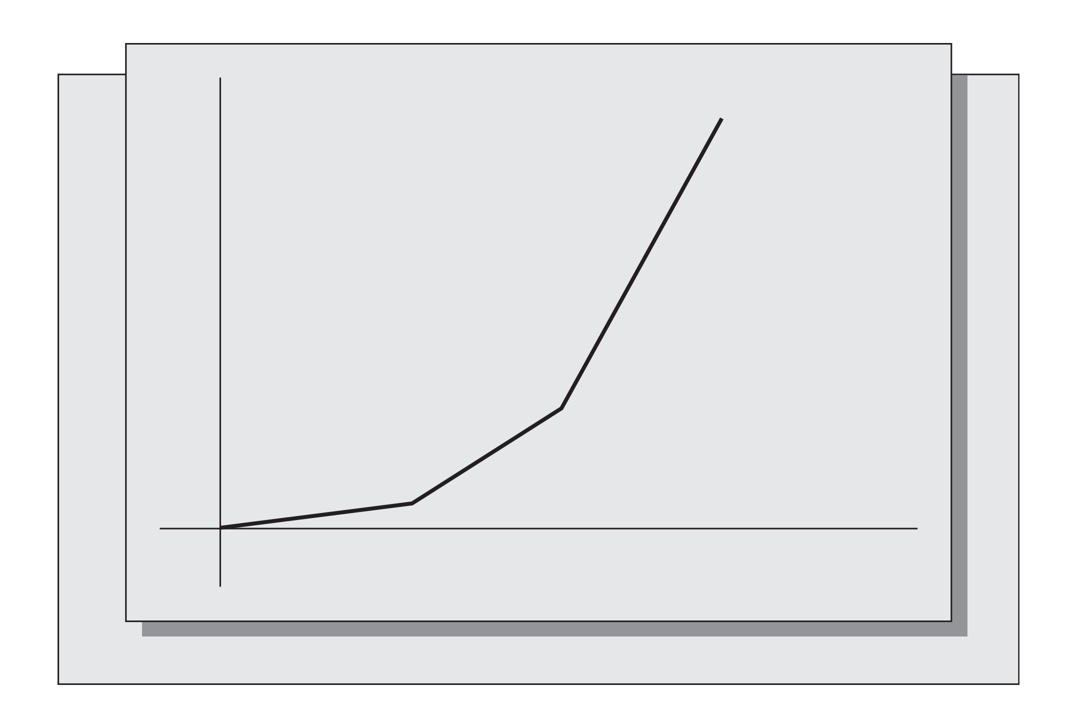
$$P_t \leq C_t \qquad \text{para } 1 \leq t \leq T.$$

La formulación de programación lineal presentada en esta sección supuso que los niveles de inventario nunca serían negativos. Sin embargo, en algunos casos podría ser deseable, o hasta necesario, permitir que la demanda supere a la oferta, por ejemplo, si la demanda pronosticada superó la capacidad de producción durante varios periodos de planeación. Para manejar la acumulación de exceso de demanda, el nivel de inventario I_t debe expresarse como la diferencia entre dos variables no negativas, digamos I_t^+ e I_t^- que satisfagan

$$I_t = I_t^+ - I_t^-,$$
$$I_t^+ \geq 0, \qquad I_t^- \geq 0.$$

El costo por mantener inventarios se cargaría ahora contra I_t^+, y el costo de penalización por pedidos atrasados (digamos c_P) contra I_t^-. Sin embargo, observe que para que la solución sea sensible, debe cumplirse que I_t^+ e I_t^- no sean ambas positivas en el mismo periodo t. Como en los casos de las variables tiempo extra y tiempo inactivo, y contratación y despido, las propiedades de la programación lineal garantizarán que esto sea válido sin tener que incluir en forma explícita la restricción $I_t^+I_t^- = 0$ en la formulación.

FIGURA 3-5
Una función lineal convexa en intervalos.



En el desarrollo del modelo de programación lineal, establecimos el requerimiento de que todas las funciones de costo deben ser lineales. Esto no es estrictamente correcto. También se puede aplicar la programación lineal cuando las funciones de costo son *funciones convexas lineales en intervalos*.

Una función convexa es aquella que tiene una pendiente cada vez mayor. Una función lineal en intervalos es aquella cuya gráfica está formada por segmentos de recta. Por lo tanto, una función convexa lineal en intervalos es la descrita por segmentos de recta cuyas pendientes son cada vez mayores. Un ejemplo característico se ilustra en la figura 3-5.

En la práctica es probable que algunas o todas las funciones costo, para planeación agregada, sean convexas. Por ejemplo, si la figura 3-5 representa el costo de contratar trabajadores, el costo marginal de contratar un trabajador adicional aumenta con la cantidad de trabajadores que ya se hayan contratado. Probablemente, esto es más exacto que suponer que el costo de contratar un trabajador más es una constante independiente de la cantidad de trabajadores contratados previamente. A medida que se contratan más trabajadores, la oferta de trabajadores se contrae, y se deben hacer más esfuerzos para contratar a los trabajadores restantes que haya disponibles.

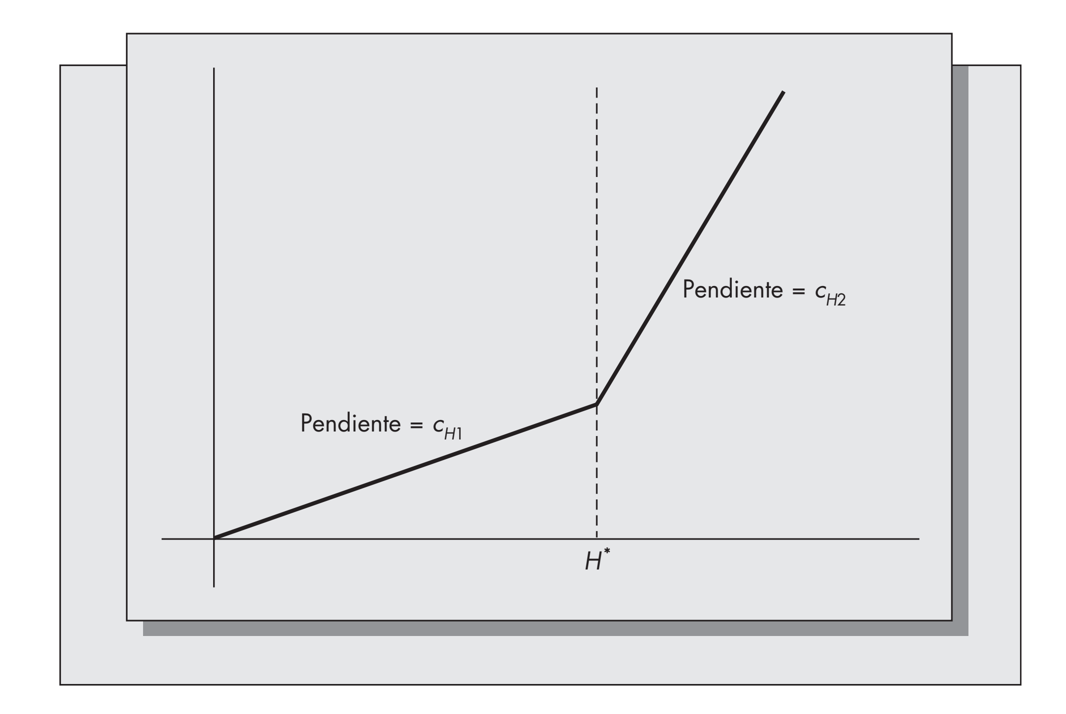
Para ver exactamente cómo se incorporarían las funciones lineales convexas en intervalos a la formulación de programación lineal, describiremos un caso muy sencillo. Supongamos que el costo de contratar nuevos trabajadores está representado por la función de la figura 3-6. De acuerdo con ésta, cuesta c_{H1} contratar cada trabajador hasta llegar a H^* trabajadores contratados, y cuesta c_{H2} contratar cada trabajador después de las H^* primeras contrataciones, con $c_{H1} < c_{H2}$. La variable H_t, o cantidad de trabajadores contratados en el periodo t, se debe expresar como la suma de dos variables:

$$H_t = H_{1t} + H_{2t}.$$

Se interpreta H_{1t} como la cantidad de trabajadores contratada hasta H^*, y H_{2t} como los contratados más allá de H^* en el periodo t. Ahora, el costo de contratar se representa en la función objetivo como

$$\sum_{t=1}^{T} (c_{H1}H_{1t} + c_{H2}H_{2t}),$$

FIGURA 3-6
Función lineal convexa en intervalos de los costos de contratación.



y las restricciones adicionales

$$H_t = H_{1t} + H_{2t}$$
$$0 \le H_{1t} \le H^*$$
$$0 \le H_{2t}$$

también deben incluirse.

Para que la solución final tenga sentido, nunca debe suceder que $H_{1t} < H^*$ y $H_{2t} > 0$ para alguna t. (¿Por qué?) Sin embargo, como la programación lineal busca la solución con costo mínimo, forzará que H_{1t} adquiera su valor máximo antes de dejar que H_{2t} sea positivo, porque $c_{H1} < c_{H2}$. Esta es la razón por la cual las funciones de costo deben ser convexas. Este método se puede ampliar con facilidad a más de dos segmentos lineales, y a cualquiera de las demás funciones de costo que haya en la función objetivo. La técnica se llama programación convexa separable, y se describe con mayor detalle en Hillier y Lieberman (1990).

Otros métodos de solución

Bowman (1956) sugirió una formulación tipo transporte del problema de planeación agregada, cuando no se permiten pedidos atrasados. Algunos autores han investigado la solución de problemas de planeación agregada con algoritmos hechos a la medida. Esos algoritmos podrían ser más rápidos para resolver problemas grandes de planeación agregada que un software general de programación lineal, como LINDO. Además, algunas de esas formulaciones permiten tener ciertos tipos de costos no lineales. Por ejemplo, si hay economías de escala en la producción, es probable que la función del costo de producción sea cóncava respecto a la cantidad de unidades producidas. Las funciones de costo cóncavas no son tan adecuadas a las formulaciones de programación lineal como las convexas. Creemos que un software de programación lineal para propósitos generales es suficiente para la mayoría de los problemas de tamaño razonable que se encuentran en el mundo real. Sin embargo, el lector debe estar consciente de que hay

técnicas alternativas de solución que podrían ser más eficientes para problemas grandes, y que podrían llegar a soluciones cuando algunas de las funciones de costo son no lineales. Un trabajo que menciona en detalle un procedimiento del tipo de transporte, más eficiente que el método simplex para resolver problemas de planeación agregada, es el de Erenguc y Tufekci (1988).

3.6 SOLUCIÓN DE PROBLEMAS DE PLANEACIÓN AGREGADA MEDIANTE PROGRAMACIÓN LINEAL: UN EJEMPLO

Demostraremos el uso de la programación lineal determinando la solución óptima del ejemplo descrito en la sección 3.4. Como no se permite que haya subcontratación, tiempo extra o tiempo inactivo, y como los coeficientes de los costos son constantes respecto al tiempo, la función objetivo es simplemente

$$\text{Minimizar}\left(500\sum_{t=1}^{6} H_t + 1\,000\sum_{t=1}^{6} F_t + 80\sum_{t=1}^{6} I_t\right).$$

Las condiciones de frontera comprenden las especificaciones del inventario inicial, de 500 unidades, la fuerza inicial de trabajo, de 300 trabajadores, y el inventario final, de 600 unidades. La mejor manera de manejarlas es incluir, por separado, una restricción adicional para cada condición de frontera.

Las restricciones se obtienen sustituyendo $t = 1, \ldots, 6$ en las ecuaciones (A), (B) y (C). El conjunto completo de restricciones expresadas en formato de programación lineal (con todas las variables del problema en el lado izquierdo y las constantes no negativas al lado derecho), es el siguiente:

$$\begin{aligned}
W_1 - W_0 - H_1 + F_1 &= 0,\\
W_2 - W_1 - H_2 + F_2 &= 0,\\
W_3 - W_2 - H_3 + F_3 &= 0,\\
W_4 - W_3 - H_4 + F_4 &= 0,\\
W_5 - W_4 - H_5 + F_5 &= 0,\\
W_6 - W_5 - H_6 + F_6 &= 0;
\end{aligned} \tag{A}$$

$$\begin{aligned}
P_1 - I_1 + I_0 &= 1\,280,\\
P_2 - I_2 + I_1 &= 640,\\
P_3 - I_3 + I_2 &= 900,\\
P_4 - I_4 + I_3 &= 1\,200,\\
P_5 - I_5 + I_4 &= 2\,000,\\
P_6 - I_6 + I_5 &= 1\,400;
\end{aligned} \tag{B}$$

$$\begin{aligned}
P_1 - 2.931W_1 &= 0,\\
P_2 - 3.517W_2 &= 0,\\
P_3 - 2.638W_3 &= 0,\\
P_4 - 3.810W_4 &= 0,\\
P_5 - 3.224W_5 &= 0,\\
P_6 - 2.198W_6 &= 0;
\end{aligned} \tag{C}$$

$$W_1, \ldots, W_6, P_1, \ldots, P_6, I_1, \ldots, I_6, F_1, \ldots, F_6, H_1, \ldots, H_6 \geq 0; \qquad \text{(D)}$$

$$W_0 = 300,$$
$$I_0 = 500, \qquad \text{(E)}$$
$$I_6 = 600.$$

Resolvimos este programa lineal usando el paquete LINDO desarrollado por Schrage (1984). Los resultados de ese programa aparecen en la tabla 3-5. En el suplemento sobre programación

TABLA 3-5 **Solución de programación lineal para el problema de Densepack**

LP OPTIMUM FOUND AT STEP 24

OBJECTIVE FUNCTION VALUE

1) 379320.900

VARIABLE	VALUE	REDUCED COST
H1	.000000	1500.000000
H2	.000000	791.378100
H3	.000000	222.439900
H4	.000000	6.736023
H5	464.782300	.000000
H6	.000000	192.750300
F1	27.047150	.000000
F2	.000000	708.621900
F3	.000000	1277.560000
F4	.000000	1493.264000
F5	.000000	1500.000000
F6	.000000	1307.250000
W0	300.000000	.000000
W1	272.952900	.000000
W2	272.952900	.000000
W3	272.952900	.000000
W4	272.952900	.000000
W5	737.735200	.000000
W6	737.735200	.000000
I0	500.000000	.000000
I1	20.024820	.000000
I2	340.000000	.000000
I3	160.049800	.000000
I4	.000000	18.445950
I5	378.458200	.000000
I6	600.000000	.000000
P1	800.024800	.000000
P2	959.975200	.000000
P3	720.049600	.000000
P4	1039.950000	.000000
P5	2378.458000	.000000
P6	1621.542000	.000000

TABLA 3-6 Plan agregado para Densepack obtenido del redondeo de la solución de programación lineal

A	B	C	D	E	F	G	H	I
Mes	**Número de trabajadores**	**Número de contratados**	**Número de despedidos**	**Número de unidades por trabajador**	**Número de unidades producidas (B × E)**	**Producción acumulada**	**Demanda neta acumulada**	**Inventario final (G − H)**
Enero	273		27	2.931	800	800	780	20
Febrero	273			3.517	960	1 760	1 420	340
Marzo	273			2.638	720	2 480	2 320	160
Abril	273			3.810	1 040	3 520	3 520	0
Mayo	738	465		3.224	2 379	5 899	5 520	379
Junio	738			2.198	1 622	7 521	7 520	1
Totales		465	27					900

lineal también describiremos la solución con Excel. Naturalmente, la función Solver de Excel da los mismos resultados. El valor de la función objetivo en la solución óptima es $379 320.90, que es considerablemente menor que la que obtuvimos con el plan de cero inventarios o el de fuerza laboral constante. Sin embargo, este costo se basa en valores fraccionarios de las variables. El costo real se elevará un poco, después de redondear.

Nos apegaremos al procedimiento de redondeo que recomendamos anteriormente para redondear todos los valores de W_t al entero inmediato superior. Con esto se obtiene $W_1 = \cdots = W_4 = 273$, y $W_5 = W_6 = 738$. Con ellos se determinan los valores de las demás variables del problema. Esto quiere decir que la empresa necesita despedir a 27 trabajadores en enero y contratar 465 en mayo. La solución completa aparece en la tabla 3-6.

Como la columna H de la tabla 3-6 corresponde a la demanda neta, sumaremos otra vez las 600 unidades del inventario final en junio, obteniendo un inventario total de 900 + 600 = 1 500 unidades. Por consiguiente, el costo total de este plan es (500)(465) + (1 000)(27) + (80)(1 500) = $379 500, lo que representa un ahorro apreciable en comparación con el plan de cero inventarios o con el plan de fuerza de trabajo constante.

Los resultados del análisis de programación lineal sugieren otro plan que podría ser más adecuado para la empresa. Como la estrategia óptima es disminuir la fuerza de trabajo en enero y reconstruirla en mayo, una alternativa razonable podría ser no despedir los 27 trabajadores en enero, y contratar menos trabajadores en mayo. En este caso, el método más eficiente para determinar la cantidad correcta de trabajadores que deben contratarse en mayo consiste en volver a resolver el programa lineal, pero sin las variables $F_1, \ldots, F_6$, puesto que el no despedir trabajadores significa que se fuerzan esas variables a ser cero. (Si el lector desea evitar la reintroducción del problema en la computadora, sólo use la anterior formulación con las restricciones $F_1 = 0, F_2 = 0, \ldots, F_6 = 0$.) La cantidad óptima de trabajadores que habrá que contratar en mayo es 374 si no se despide a ninguno en enero, y el costo aproximado del plan es 386 120 dólares. Cuesta un poco más que el plan óptimo, y tiene la importante ventaja de no requerir el despido de ninguno de los trabajadores.

Problemas para las secciones 3.5 y 3.6

17. Formule el problema 13 como un programa lineal. Asegúrese de definir todas las variables e incluya todas las restricciones requeridas.

18. El problema 17 se resolvió en la computadora. El resultado de la programación lineal es el siguiente:

```
LP OPTIMUM FOUND AT STEP        14

                OBJECTIVE FUNCTION VALUE

 1)      36185.0600

VARIABLE            VALUE             REDUCED COST
      H1          6.143851               .000000
      H2         64.310690               .000000
      H3           .000000            300.000000
      H4        116.071400               .000000
      F1           .000000            300.000000
      F2           .000000            300.000000
      F3         87.662330               .000000
      F4           .000000            300.000000
      I1           .000000             59.415580
      I2           .000000            189.285700
      I3           .000000             31.006490
      W1        106.143900               .000000
      W2        170.454500               .000000
      W3         82.792210               .000000
      W4        198.863600               .000000
      P1        850.000000               .000000
      P2       1260.000000               .000000
      P3        510.000000               .000000
      P4        980.000000               .000000
      I4           .000000            120.292200
```

a) Determine una solución factible para el problema, redondeando los valores de las variables mediante el método conservador descrito en la sección 3.5. Calcule el costo del plan resultante.

b) Determine otro plan que logre un costo menor que el calculado en el inciso *a*), manteniendo la factibilidad (esto es, que tenga una cantidad no negativa de inventario disponible al final de cada periodo).

19. *a*) Formule el problema 9 como un programa lineal.

b) Resuelva el problema con un software de programación lineal. Redondee las variables en la solución y determine el costo del plan obtenido.

*20. *a*) Formule el problema 14 como un programa lineal.

b) Resuelva el problema con un paquete de programación lineal. Redondee la solución y determine el costo del plan resultante.

21. Considere a Mr. Meadows Cookie Company del problema 13. Suponga que el costo de contratar trabajadores cada periodo es $100 por cada uno hasta llegar a 20 trabajadores, $400 por cada uno de 21 a 50, y $700 por cada uno contratado más allá de 50.

a) Escriba la formulación completa de programación lineal para este problema modificado.

b) Resuelva el problema modificado con un paquete de programación lineal. ¿Qué diferencia produce la nueva función de costo de contratación en la solución?

*20. Leather-All, produce una línea de productos de piel hechos a mano. En la actualidad sólo produce cinturones, bolsas de mano y portafolios. La demanda predicha para esos tres artículos, en un horizonte de planeación de seis meses, es la siguiente:

Mes	Días hábiles	Cinturones	Bolsas	Portafolios
1	22	2 500	1 250	240
2	20	2 800	680	380
3	19	2 000	1 625	110
4	24	3 400	745	75
5	21	3 000	835	126
6	17	1 600	375	45

La producción de un cinturón requiere dos horas, una bolsa tres horas y un portafolio seis horas. Todos los trabajadores tienen la destreza suficiente para trabajar en cualquier artículo. Leather-All tiene 46 empleados con acciones de la empresa y no se pueden despedir. Hay otros 30 trabajadores locales disponibles, que se pueden contratar por cortos periodos a mayor costo. Los empleados regulares ganan 8.50 dólares por hora en tiempo normal, y 14 dólares por hora en tiempo extra. El tiempo regular consiste en días de trabajo de siete horas, y los empleados regulares pueden trabajar todo el tiempo extra que haya disponible.

Los trabajadores adicionales se contratan a 11 dólares por hora, y se mantienen en la nómina un mínimo de un mes completo. Los costos de contratación y despido son despreciables.

Debido a la naturaleza competitiva del ramo, Leather-All no desea incurrir en pedidos surtidos con retraso.

a) Use las horas-hombre como medida agregada de la producción y convierta las demandas pronosticadas a demandas en términos de unidades agregadas.

b) ¿Cuál sería el tamaño de la fuerza de trabajo necesaria para satisfacer la demanda de los seis meses próximos trabajando sólo tiempo normal? ¿Sería ventajoso para la empresa llevar a la fuerza de trabajo permanente hasta este nivel? ¿Por qué sí o por qué no?

c) Determine un plan de producción que cumpla con la demanda usando sólo empleados de tiempo normal, y el costo total de ese plan.

d) Determine un plan de producción que sólo use empleados adicionales para absorber el exceso de demanda, y calcule el costo del plan.

*23. *a*) Formule el problema de optimizar el programa de contrataciones de Leather-All como un programa lineal. Defina todas las variables del problema e incluya todas las restricciones que sean necesarias.

b) Resuelva el problema formulado en el inciso *a*) usando un software de programación lineal. Redondee todas las variables relevantes y determine el costo del plan resultante. Compare sus resultados con los que obtuvo en el problema 22*c*) y *d*).

3.7 LA REGLA LINEAL DE DECISIÓN

Holt, Modigliani, Muth y Simon (1960) sugirieron un método alternativo interesante para resolver el problema de planeación agregada. Suponen que todos los costos relevantes, incluyendo los de inventario y los de cambiar niveles de producción y cantidades de trabajadores, se representan con funciones cuadráticas. Esto es, que el costo total durante el horizonte de planeación de T periodos se puede escribir en la forma

$$\sum_{t=1}^{T} [c_1 W_t + c_2(W_t - W_{t-1})^2 + c_3(P_t - Kn_t W_t)^2 + c_4 P_t + c_5(I_t - c_6)^2]$$

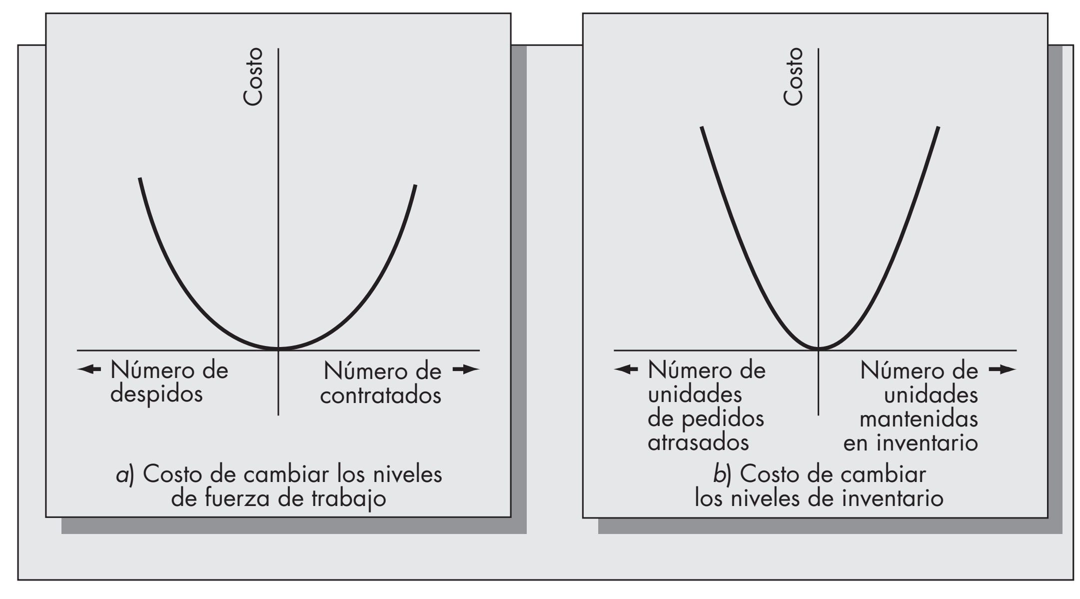
sujeta a

$$I_t = I_{t-1} + P_t - D_t \qquad \text{para } 1 \leq t \leq T.$$

Se deben determinar los valores de las constantes $c_1, c_2, \ldots, c_6$ para cada aplicación particular. Al expresar las funciones de costo como cuadráticas y no como simples funciones lineales se obtienen ciertas ventajas sobre el método de programación lineal. Como las funciones cuadráticas son diferenciables, pueden usarse las reglas normales del cálculo para determinar soluciones óptimas. La solución óptima se presentará donde las primeras derivadas parciales con respecto a cada una de las variables del problema sean iguales a cero. Cuando se diferencian las funciones cuadráticas producen ecuaciones de primer orden (ecuaciones lineales), fáciles de resolver. También las funciones cuadráticas proporcionan una mejor aproximación a las funciones generales no lineales que las funciones lineales. (Toda función no lineal se puede aproximar con un desarrollo en serie de Taylor, en el que los dos primeros términos producen una aproximación cuadrática.)

Sin embargo, el enfoque cuadrático también tiene una desventaja grave: que las funciones cuadráticas (esto es, las parábolas) son simétricas (véase la figura 3-7). Por lo tanto, el costo de

FIGURA 3-7
Funciones de costo cuadrático utilizadas para la obtención de la regla de decisión lineal.



contratar determinada cantidad de trabajadores debe ser igual que el de despedir la misma cantidad, y el costo de producir determinada cantidad de unidades en tiempo extra debe ser igual que el costo asignado a la misma cantidad de unidades de tiempo inactivo de trabajador. El problema puede superarse en cierta forma si no se establece el centro de simetría de la función costo a cero, pero el problema básico de la simetría sigue existiendo.

La particularidad más atractiva de este método es la sencillez de la política óptima. Por ejemplo, P_t, el nivel óptimo de producción en el periodo t, tiene la forma

$$P_t = \sum_{n=0}^{K} (a_n D_{t+n} + bW_{t-1} + cI_{t-1} + d).$$

Los términos a_n, b, c y d son constantes que dependen de los parámetros de costo. W_t tiene una forma similar. Las ventajas computacionales de la regla lineal de decisión son menos importantes ahora que cuando se propuso por primera vez este análisis, por la disponibilidad general de computadoras hoy en día. Como la programación lineal suministra un marco mucho más flexible para formular problemas de planeación agregada y se pueden resolver programas razonablemente grandes aun con computadoras personales, podría llegarse a la conclusión de que la programación lineal es preferible como técnica de solución.

Sin embargo, hay que reconocer que la publicación de Holt, Modigliani, Muth y Simon representa un hito en la aplicación de métodos cuantitativos a problemas de planeación de la producción. Los autores desarrollaron un método de solución que da como resultado un conjunto de fórmulas fáciles de implementar, y en realidad acometieron la implementación del método. Ese trabajo detalla la aplicación del método con un gran fabricante de pinturas domésticas en el área de Pittsburgh. En realidad, el análisis fue implementado en esa compañía, pero una visita posterior a la empresa indicó que surgieron problemas serios cuando se adoptó la regla lineal de decisión, principalmente debido a la política de la firma de no despedir trabajadores cuando el modelo indicaba que eso era lo que tenía que hacerse. (Véase Vollman, Berry y Whybark, 1992, p. 627.)

3.8 MODELADO DEL COMPORTAMIENTO GERENCIAL

Bowman (1983) desarrolló una técnica interesante de planeación agregada que debería poder aplicarse también a problemas de otros tipos. Su idea fue construir un modelo sensible para controlar los niveles de producción y ajustar sus parámetros tanto como sea posible a las decisiones previas reales de la gerencia. De esta forma el modelo refleja el juicio y la experiencia de los gerentes, y evita los problemas que surgen al usar los métodos tradicionales de modelado. Uno de esos problemas es determinar la precisión de los supuestos que requiere el modelo. Otro problema que se evita es la necesidad de determinar valores de parámetros de entrada que pudieran ser difíciles de medir.

Consideremos el problema de producir un solo producto durante T periodos de planeación. Suponga que $D_1, \ldots, D_T$ son los pronósticos de la demanda durante los siguientes T periodos, y que se deben determinar los niveles de producción $P_1, \ldots, P_T$. La regla de decisión razonable más sencilla es

$$P_t = D_t \qquad \text{para } 1 \le t \le T.$$

Al tratar de hacer coincidir exactamente la producción con la demanda, por lo general se obtendrá un programa de producción demasiado errático. Se puede lograr un suavizamiento de la producción con una regla de decisión de la forma

$$P_t = D_t + \alpha(P_{t-1} - D_t),$$

donde α es un factor de suavizamiento para la producción, el cual tiene la propiedad de que $0 \leq \alpha \leq 1$. (El efecto del suavizamiento, en este caso, es igual que en el suavizamiento exponencial, descrito con detalle en el capítulo 2.) Cuando α es igual a cero, esta regla es idéntica a la anterior. Cuando $\alpha = 1$, la producción en el periodo t es exactamente igual a la del periodo $t - 1$. La elección de α permite asignar un peso relativo a la coincidencia de la producción con la demanda, en comparación con mantener constante la producción de un periodo al siguiente.

Además de suavizar la producción, la empresa también podría tratar de mantener niveles de inventario cercanos a cierto nivel predeterminado, por ejemplo I_N. Un modelo que suaviza en forma simultánea los inventarios y la producción es

$$P_t = D_t + \alpha(P_{t-1} - D_t) + \beta(I_N - I_{t-1}),$$

donde $0 \leq \beta \leq 1$ mide el peso relativo asignado al suavizamiento del inventario.

Por último, el modelo debe incorporar pronósticos de demanda. De esta forma se pueden aumentar o disminuir los niveles de producción en forma anticipada a un cambio en el patrón de las demandas. Una regla que incluye el suavizamiento de la producción, del inventario y de los pronósticos de la demanda en el futuro es

$$P_t = \sum_{i=t}^{t+n} a_{t-i+1} D_i + \alpha(P_{t-1} - D_t) + \beta(I_N - I_{t-1}).$$

A esta regla de decisión se llega a través de un análisis directo y con sentido común de lo que debe ser una buena regla de producción. Es interesante notar lo parecida que es esta regla a la de decisión lineal descrita en la sección anterior, que se dedujo de un modelo matemático. Sin duda, el trabajo previo sobre la regla lineal de decisión inspiró la forma de esta regla. Con frecuencia se da el caso de que la propiedad más valiosa de un modelo es indicar la forma óptima de una estrategia óptima, y no necesariamente la estrategia óptima misma.

Este modelo particular de P_1 requiere determinar $a_1, \ldots, a_n$, α, β e I_N. Bowman sugiere determinar los valores de estos parámetros en forma retrospectiva, observando al sistema durante un intervalo razonable y usando una técnica como la de los mínimos cuadrados. De esta forma, el modelo se transforma en una reflexión del comportamiento gerencial en el pasado.

Bowman compara la experiencia real de varias compañías con aquellas que podrían haber usado este método, y demuestra que en la mayoría de los casos pudieran obtenerse reducciones apreciables de costo. Si el modelo sólo emula el comportamiento gerencial, ¿por qué mejoraría entonces la experiencia administrativa real? La teoría es que un modelo deducido en esta forma es una reflexión de la gerencia al tomar decisiones racionales, ya que casi siempre el sistema resulta estable. Sin embargo, si se presenta un evento repentino y desacostumbrado, como una demanda mucho mayor que la anticipada, o una disminución repentina de la capacidad productiva debida, por ejemplo, a la descompostura de una máquina o a la pérdida de personal clave, es probable que muchos gerentes tiendan a reaccionar precipitadamente. Sin embargo, el modelo recomendaría niveles de producción que sean consistentes con la toma de decisiones en el pasado. Por consiguiente, emplear un modelo simple pero consistente de la toma de decisiones evitaría que la gerencia sintiera pánico en una situación desacostumbrada. Aunque el método es conceptualmente atractivo, hay pocos informes publicados de buenas implementaciones de esta técnica.

Problemas para las secciones 3.7 y 3.8

24. La forma de la regla de producción óptima deducida por Holt, Modigliani, Muth y Simon (1960) para el caso de la empresa de pinturas indica que el nivel de producción en el periodo t debe calcularse con la fórmula

$$\begin{aligned} P_t = {} & .463D_t + .234D_{t+1} + .111D_{t+2} + .046D_{t+3} + .013D_{t+4} \\ & - .022D_{t+5} - .008D_{t+6} - .010D_{t+7} - .009D_{t+8} \\ & - .008D_{t+9} - .007D_{t+10} - .005D_{t+11} \\ & + .993W_{t-1} - .464I_{t-1} + 153, \end{aligned}$$

y que el nivel de fuerza de trabajo en el periodo t se calcule con la fórmula

$$\begin{aligned} W_t = {} & .010D_t + .0088D_{t+1} + .0071D_{t+2} + .0054D_{t+3} + .0042D_{t+4} \\ & + .0031D_{t+5} + .0023D_{t+6} + .0016D_{t+7} + .0012D_{t+8} \\ & + .0009D_{t+9} + .0006D_{t+10} + .0005D_{t+11} \\ & + .743W_{t-1} - .01I_{t-1} - 2.09. \end{aligned}$$

Suponga que las demandas pronosticadas para los 12 meses siguientes son 150, 164, 185, 193, 167, 143, 126, 93, 84, 72, 50 y 66. La cantidad actual de trabajadores es 180, y el nivel actual de inventarios es 45 unidades agregadas.

a) Calcule los valores del nivel de producción agregada y el número de empleados que debe usar la empresa en el periodo actual.

b) Repita el cálculo del inciso *a*), pero no tome en cuenta todos los términos cuyo coeficiente tenga un valor absoluto menor o igual que .01, al calcular P_t, y menor o igual que .005 al calcular W_t. ¿Cuánta diferencia observa en las respuestas?

c) Suponga que la demanda actual se evalúa y resulta 163, y que el nuevo pronóstico a un año es 72. Usando la aproximación sugerida en el inciso *b*), determine los nuevos valores para el nivel de producción agregada y la cantidad de trabajadores. [Sugerencia: Debe calcular el nuevo valor de I_{t-1} y suponer que el nuevo valor de W_{t-1} es el que calculó en el inciso *a*).]

25. Use los siguientes valores de coeficientes gerenciales para el modelo de Bowman, de producción suavizada, para determinar el nivel de producción que debe planear Harold Grey (descrito en el problema 9) para el año siguiente. Suponga que el nivel actual de producción es 150 000 paquetes.

$$a_1 = .3475, \quad a_2 = .1211, \quad a_3 = .0556, \quad a_4 = .0663,$$
$$a_5 = .0023, \quad \alpha = .6, \quad \beta = .3, \quad I_N = 40\,000.$$

26. *a*) Describa las semejanzas y diferencias entre la regla lineal de decisión de Holt, Modigliani, Muth y Simon, y el método de coeficientes gerenciales de Bowman.

b) Describa las ventajas y desventajas relativas de las reglas lineales de decisión contra las de programación lineal para resolver problemas de planeación agregada.

3.9 DESAGREGACIÓN DE PLANES AGREGADOS

Se puede hacer planeación agregada a varios niveles, pero siempre se da el caso de que un plan agregado corresponde a cierto agrupamiento de artículos. Los agrupamientos pueden variar de tamaño, desde algunos que abarcan una familia de productos hasta toda la línea de artículos que

produce la empresa. En cualquier caso, es importante asegurarse de que el plan agregado y los programas maestros de producción para artículos individuales sean consistentes. Cuando los planes agregados resultan de combinar o agregar planes de artículos individuales, la desagregación es directa. Sin embargo, la planeación agregada proporciona una perspectiva administrativa de arriba abajo, y por esa razón los planes agregados no son, por lo general, el resultado de una agregación de abajo hacia arriba de programas maestros de producción de artículos individuales. Una empresa que trate de obtener consistencia entre el plan agregado y los programas de producción de artículos individuales necesita desarrollar un esquema de desagregación consistente con la definición de una unidad agregada y con la estructura organizacional de la compañía.

No es muy evidente que la coordinación de los programas de producción para artículos individuales con el plan agregado sea un asunto serio en todos los casos. Se puede considerar que la planeación agregada es principalmente un medio de calcular los niveles aproximados de fuerza de trabajo durante determinado horizonte de planeación. Las unidades agregadas pueden ser un ente ficticio cuyo único propósito es desarrollar una representación lógica de la actividad general de ventas de la empresa. Cuando este es el caso, la planeación detallada de la producción de artículos individuales se puede considerar como un problema aparte que debe resolverse con apego a las restricciones determinadas por el tamaño de la fuerza de trabajo.

Sin embargo, si una unidad de producción agregada sí corresponde a un artículo real, o se determina mediante una estructura jerárquica de planeación como la que ya se describió en el capítulo (artículos, familias y tipos), es importante que el programa maestro de producción y el plan agregado sean consistentes. El problema de descomponer un plan agregado en un programa maestro por artículo ha recibido escasa atención. Es probable que las empresas individuales estén manejando el problema en formas que se adecuen a sus sistemas específicos de producción y a sus estructuras organizacionales, pero se han desarrollado pocos métodos generales para resolver el problema.

Un esquema general para manejar el asunto de la desagregación, que es consistente con el agrupamiento de artículos en familias y de familias en tipos, ha sido promovido por un grupo de investigadores del MIT. Sus métodos se describen con detalle en la publicación de Hay y Candea (1984). Presentaremos una perspectiva muy breve de su método.

Suponga que X^* representa el número de unidades agregadas de producción de determinado tipo, indicado por el plan agregado para el periodo de planeación inmediato siguiente. Se calcula la cantidad de unidades, Y, que se producirán de la familia j, resolviendo el siguiente problema matemático de programación:

$$\text{Minimizar} \sum_{j=1}^{J} \frac{K_j \lambda_j}{Y_j}$$

sujeta a

$$\sum_{j=1}^{J} Y_j = X^*$$

y

$$a_j \leq Y_j \leq b_j \qquad \text{para } 1 \leq j \leq J,$$

donde K_j es el costo de preparación para la familia j y λ_j es la demanda anual para la familia j. Las constantes a_j y b_j son cotas superior e inferior de la cantidad de unidades que pueden producirse por cada familia. El término $K_j\lambda_j/Y_j$ es el costo promedio anual para la familia j.

Aplicación instantánea

WELCH'S USA LA PLANEACIÓN AGREGADA PARA PROGRAMAR SU PRODUCCIÓN

Welch's es un fabricante de alimentos preparados para su almacenamiento, con base en Concord, Massachusetts. Welch's es conocido por su jalea y jugo de uva, pero también produce una gran variedad de alimentos procesados. Allen y Schuster (1994) describen un modelo de la planeación agregada que se implementa en la fábrica más grande de Welch's.

Las características del sistema de producción para el que se diseñó la planeación son:

- Demanda dinámica, incierta, que tiene como consecuencia el cambio en los requerimientos de inventarios de reserva.
- Ambiente tipo "fabricar para almacenar".
- Líneas de producción dedicadas.
- Líneas de producción para dos o más familias de productos cada una.
- Tiempos y costos de preparación grandes para las familias, y pequeños para los artículos individuales.

Los dos objetivos principales del sistema de producción, según los autores, son suavizar los picos de demanda a través del tiempo, para no exceder la capacidad de producción y asignar los requerimientos productivos entre las familias para equilibrar los costos de inventario y de preparación por familia. La planeación se hace con un horizonte de tiempo de seis meses para pronosticar la demanda. Este horizonte se divide en dos: las cuatro semanas siguientes y los cinco meses restantes. Se preparan planes detallados para el corto plazo, incluyendo asignación de producción normal y en tiempo extra.

El modelo tiene dos componentes primarios: un modelo de planeación por familia, que determina la programación y el dimensionamiento de las corridas de producción por familia, y un modelo de desagregación de planeación, que toma los resultados de la planeación por familia y determina los tamaños de lote para artículos individuales dentro de las familias.

Los autores describen también varios aspectos de la implementación, específicos para el ambiente de Welch's. Las duraciones de las corridas de producción se deben ligar a la estructura actual de turnos de ocho horas. Para hacerlo, recomiendan que las longitudes de corridas de producción se expresen como múltiplos de un cuarto de turno (dos horas).

El modelo fue implementado en una computadora personal. Los tiempos de cómputo son muy moderados. Entre las técnicas de solución se encuentran un programa matemático entero mixto y una formulación de programación lineal con relajación (esto es, redondeo de las variables a valores enteros).

Este caso demuestra que los conceptos descritos en este capítulo pueden ser de utilidad en un ambiente real de planeación de producción. Aunque el sistema aquí descrito no se basa en ninguno de los modelos específicos mencionados en este capítulo, esta aplicación demuestra que la agregación y la desagregación son conceptos útiles. La desagregación jerárquica para programación de producción es una valiosa herramienta de planeación.

(En el capítulo 4 se presenta una deducción detallada de este término, al explicar modelos para administración de artículos individuales.) Por consiguiente, la función objetivo representa el costo ponderado anual total de preparación para todas las familias que forman este tipo particular. En este nivel no se toma en cuenta el costo por mantener inventarios porque se supone que la función objetivo para el modelo de planeación agregada que dio como resultado X^* ya incluye dicho costo. La justificación de usar sólo los resultados del primer periodo del plan agregado es que los programas se generan en una base de horizonte progresivo. A medida que se dispone de nueva información en cada periodo, se corrigen y actualizan los programas anteriores.

Es necesario considerar varios aspectos de la factibilidad antes de desagregar más los tamaños de corrida para familias y obtener lotes para artículos individuales. Lo que se pretende es programar los lotes de artículos individuales dentro de una familia, de tal manera que se agoten en el tiempo programado de preparación para la familia. De esta forma, se pueden producir los artículos de la misma familia con la misma preparación de producción.

El concepto de desagregar el plan agregado siguiendo las líneas organizacionales de manera que sea consistente con el esquema de agregación es muy atractivo. Queda por ver si los métodos descritos en esta sección proporcionan un enlace funcional entre los planes agregados y los programas detallados por artículo.

Chung y Krajewski (1984) han investigado otro método para atacar el problema de la desagregación. Desarrollan una formulación matemática de programación para el problema. Los datos del programa incluyen planes agregados para cada familia de productos. Para esto se incluyen tiempo de preparación, estado de preparación, nivel total de producción para la familia, nivel de inventarios, nivel de fuerza de trabajo y disponibilidad de tiempo extra o de tiempo regular. El objetivo del análisis es especificar tamaños de lote y programación de las corridas de producción para cada artículo individual, consistentes con la información agregada para la familia de productos. Aunque esa información proporciona un eslabón potencial entre el plan agregado y el programa maestro de producción, el programa matemático que resulta requiere muchos datos y puede dar como resultado un problema muy grande que pudiera tardar en resolverse. Por esas razones no es probable que las empresas usen esos métodos cuando hay grandes cantidades de artículos individuales en cada familia de productos.

Problemas para la sección 3.9

27. ¿Qué significa "desagregación de planes agregados"?
28. Comente la siguiente cita, atribuida a un gerente de producción: "La planeación agregada es inútil para mí, porque los resultados no tienen nada que ver con mi programa maestro de producción".

3.10 PLANEACIÓN DE LA PRODUCCIÓN A ESCALA GLOBAL

La globalización de las operaciones de manufactura ya es común. Muchas grandes corporaciones se clasifican hoy como multinacionales; las actividades de manufactura y distribución cruzan las fronteras internacionales. Con la globalización de las fuentes de producción y de los mercados, las empresas deben replantear las estrategias de planeación de la producción. Un tema que se investigó en este capítulo fue el suavizamiento de los planes de producción a través del tiempo; los costos de aumentar o disminuir los niveles de fuerza de trabajo (y en consecuencia, los niveles de producción) desempeñan un papel principal en la optimización de cualquier plan agregado. Al formular las estrategias globales de producción surgen otros aspectos del suavizamiento. Las tasas de cambio, los costos de la mano de obra directa y la estructura impositiva son sólo algunas de las diferencias entre países que deben factorizarse dentro de una estrategia global.

¿A qué se debe el creciente interés en las operaciones globales? Respondiendo en breve: al costo y a la competitividad. Según McGrath y Bequillard (1989):

> Los beneficios de una estrategia manufacturera internacional bien ejecutada pueden ser muy sustanciales. Una estrategia bien desarrollada puede tener un impacto directo en el desempeño financiero, y en última instancia se reflejará en mayor rentabilidad. En la industria electrónica hay ejemplos de empresas que obtienen de 5 a 15% de reducción de costos de los bienes que venden, incrementos de 10 a 25% en las ventas, mejora de 50 a 150% en la utilización de activos y de 30 a 100% de aumento en la rotación de inventarios, y lo atribuyen a la internacionalización de la manufactura.

Cohen y otros (1989) mencionan algunos aspectos que deben caracterizar a una empresa que planea obtener niveles de producción en base mundial. Entre éstos se incluyen los siguientes:

- Para lograr los tipos de economía de escala requeridos y ser competitivos en la actualidad, las plantas y los vendedores multinacionales deben administrarse como un sistema global.
- Los aranceles y las tarifas se basan en flujos de materiales. Su impacto debe entrar en las decisiones acerca de embarques de materias primas, productos intermedios y productos terminados a través de las fronteras nacionales.

- Las tasas de cambio fluctúan al azar y afectan a los costos de producción y a las decisiones de precios en países donde el producto se produce y se vende.
- Los impuestos corporativos varían mucho de un país a otro.
- El abastecimiento global debe tener en cuenta mayores tiempos de demora, menores costos unitarios y acceso a nuevas tecnologías.
- Las estrategias de penetración de mercado, reglas de contenido local y cuotas restringen el flujo por las fronteras.
- Los diseños del producto pueden variar de acuerdo con el mercado nacional.
- El control centralizado de empresas multinacionales crea dificultades por varias razones, y el control descentralizado requiere de coordinación.
- Las diferencias culturales, de lenguaje y de las destrezas pueden ser importantes.

La determinación de las estrategias óptimas globalizadas de manufactura es un problema intimidante para cualquier firma internacional. Se pueden formular y resolver problemas matemáticos semejantes a las formulaciones de programación lineal de los modelos de planeación agregada presentados en este capítulo, pero siempre deben compararse los resultados de esos modelos con juicio y experiencia. Cohen y otros (1989) consideran uno de esos modelos. Suponen múltiples productos, plantas, mercados, materias primas, vendedores, alternativas de contratos con proveedores, periodos y países. Su formulación es un programa no mezclado, no lineal, entero, mixto y de gran escala.

Un aspecto que no se considera en su modelo es el de las fluctuaciones del tipo de cambio monetario y su efecto tanto en el precio como en la planeación de la producción. En especial, la asignación del precio se hace usualmente sumando un incremento a los costos unitarios en el mercado original. Con eso se ignora por completo el asunto de las fluctuaciones del tipo de cambio, y se puede llegar a precios irrazonables en algunos países. Por ejemplo, este asunto se presentó en Caterpillar Tractor (Caterpillar Tractor Company, 1985). En este caso, a los vendedores de todo el mundo se les proporcionaron listas de precios en dólares estadounidenses, con base en los costos de producción en ese país. Cuando el dólar era sólido en relación con otras monedas, los precios de menudeo a los clientes en el extranjero no eran competitivos en sus mercados locales. Caterpillar encontró que como resultado había perdido parte del mercado en el extranjero. A principios de los ochenta la empresa cambió a una estrategia de precios localmente competitiva para contrarrestar este problema.

El concepto de que la capacidad de manufactura se puede usar como escudo contra fluctuaciones de la tasa de cambio ha sido investigado por varios autores. Por ejemplo, Kogut y Kulatilaka (1994) desarrollan un modelo matemático para determinar cuándo es óptimo cambiar de lugar la producción. Como el costo de cambiar se supone positivo, debe haber una diferencia suficientemente grande en los tipos de cambio antes de recomendar el cambio. Por ejemplo, describen un caso en el que una empresa puede fabricar su producto tanto en Estados Unidos como en Alemania. Si en la actualidad la producción se hace en uno de los países, el modelo proporciona un medio para determinar si es económico cambiar, con base en las relativas fuerzas del marco alemán y del dólar. Si bien esos modelos se encuentran en sus primeras etapas de desarrollo, permiten contar con un medio para racionalizar las estrategias de planeación de producción internacional. Asuntos similares también han sido investigados por Huchzermeier y Cohen (1996).

3.11 CONSIDERACIONES PRÁCTICAS

La planeación agregada puede ser una ayuda valiosa para planear los niveles de producción y de mano de obra en una compañía, y es un medio de absorber las fluctuaciones de la demanda al suavizar los niveles de producción y fuerza de trabajo. Hay varias ventajas de la planeación a nivel de agregado en comparación con el nivel de detalle. Una es que el costo de preparar

pronósticos y determinar los parámetros de productividad y costo con base en un producto individual puede ser prohibitivo. El costo de recopilación y captura de datos es, probablemente, una desventaja muy importante en las formulaciones matemáticas, a gran escala de planes detallados de producción; recopilar y capturar datos cuesta más que realizar los cálculos. Una segunda ventaja de la planeación agregada es la mejora relativa de la precisión de pronóstico que se puede lograr agregando los artículos. Como se vio en el capítulo 2, los pronósticos agregados son, por lo general, más exactos que los pronósticos individuales.[3] Por último, un marco de planeación agregada permite al administrador ver la "panorámica" y no ser influido demasiado por los detalles.

Con todas esas ventajas se esperaría que la planeación agregada tuviera un lugar importante en la planeación de las actividades productivas de casi todas las empresas. Sin embargo, al parecer no es así. Existen varios motivos para la falta de interés en la metodología de la planeación agregada. En primer lugar figura la dificultad para definir en forma correcta una unidad agregada de producción. Parece que no hay una forma sencilla de especificar cómo deben agregarse los artículos individuales que funcione en todos los casos. En segundo lugar, una vez desarrollado un esquema de agregación, se requieren estimados de costo y de demanda para las unidades agregadas. Es bastante difícil obtener información exacta sobre costo y demanda en unidades reales. Obtener esa información sobre una base agregada podría ser mucho más difícil, dependiendo del esquema de agregación que se use. En tercer lugar, los modelos de planeación agregada casi nunca reflejan las realidades políticas y operacionales del ambiente en el que trabaja la empresa. Probablemente no es muy realista suponer que se puedan cambiar con facilidad los niveles de fuerza de trabajo en casi todas las empresas. Por último, como sugieren Silver y Peterson (1985), los administradores no desean confiar en un modelo matemático para obtener respuestas a los asuntos extremadamente importantes que se manejan en este análisis.

Otro aspecto es si pueden lograrse los objetivos del análisis de planeación agregada a través de métodos que no sean los descritos en este capítulo. Schwarz y Johnson (1978) señalan que los ahorros de costo logrados con una regla lineal de decisión para planeación agregada podrían obtenerse con sólo administrar mejor el inventario agregado. Basan su hipótesis en los datos de Holt, Modigliani, Muth y Simon (1960), y demuestran que para el caso del fabricante de pintura casi todos los ahorros de costo logrados con la regla lineal de decisión fueron resultado de aumentar el inventario de reserva y no de grandes cambios en el tamaño de la fuerza laboral. Esta sola comparación no comprueba la hipótesis del autor en general, pero sugiere que una mejor administración de inventarios, con los métodos que se describirán en los capítulos 4 y 5, podría recuperar una gran parte de los beneficios que puede obtener una empresa con la planeación agregada.

Aun con esta larga lista de desventajas, los modelos matemáticos descritos en este capítulo pueden y deben servir como auxiliares a los planeadores de producción. Aunque las soluciones óptimas para un modelo matemático pueden no ser las idóneas para el problema que se trata de resolver, sí proporcionan una perspectiva y pueden revelar alternativas que de otra forma no serían evidentes.

3.12 NOTAS HISTÓRICAS

El problema de planeación agregada fue concebido en una importante serie de trabajos publicados a mitad de los años cincuenta. El primero, de Holt, Modigliani y Simon (1955), describía la estructura del problema e introducía el método de costos cuadráticos, y el estudio posterior, de Holt, Modigliani y Muth (1956), se concentró en los aspectos computacionales del

[3] Con más precisión, si los pronósticos individuales no se correlacionan perfectamente, la desviación estándar del error de pronóstico será menor que la suma de las desviaciones estándar de los pronósticos individuales.

modelo. Una descripción completa del método y su aplicación a la planeación de producción para un fabricante de pinturas fue presentada por Holt, Modigliani, Muth y Simon (1960).

Parece que a principios de la década de los cincuenta ya se sabía que los problemas de planeación de producción podían formularse como programas lineales. Bowman (1956) describió el uso de un modelo de transporte para la planeación de producción. La formulación particular de programación lineal del problema de planeación agregada, descrita en la sección 3.5, es esencialmente igual a la desarrollada por Hansmann y Hess (1960). Otras formulaciones con programación lineal de ese problema implican, por lo general, productos múltiples o estructuras de costo más complejas (véase, por ejemplo, Newson, 1975a y 1975b).

Los trabajos más recientes sobre el problema de la planeación agregada se han concentrado en asuntos de agregación y desagregación (Axsater, 1981; Bitran y Hax, 1981; y Zoller, 1971), en la incorporación de curvas de aprendizaje a las reglas lineales de decisión (Ebert, 1976), en extensiones para tener en cuenta varios productos (Bergstrom y Smith, 1970) y en la inclusión de variables mercadotécnicas y/o financieras (Damon y Schramm, 1972; Leitch, 1974).

Taubert (1968) considera una técnica a la que llama regla de decisión de búsqueda. Ese método requiere desarrollar un modelo computarizado de simulación del sistema, e investigar la superficie de respuesta usando técnicas normales de búsqueda para obtener una solución (no necesariamente óptima). El método de Taubert, explicado con detalle por Buffa y Taubert (1972), da resultados comparables con los de Holt, Modigliani, Muth y Simon (1960) para el caso de la empresa fabricante de pinturas.

Kamien y Li (1990) desarrollaron un modelo matemático para examinar los efectos de subcontratar sobre decisiones de planeación agregada. Los autores demuestran que, bajo ciertas circunstancias, es preferible producir en la empresa, y presenta métodos adicionales para suavizar los niveles de producción y de fuerza de trabajo.

3.13 Resumen

La determinación de los niveles óptimos de producción para todos los productos que elabora una empresa grande puede ser una enorme tarea. La *planeación agregada* ataca este problema suponiendo que pueden agruparse artículos individuales. Sin embargo, puede ser difícil encontrar un esquema de agregación efectivo. El esquema que se elija debe ser consistente con la estructura de la organización y con la naturaleza de los productos que se fabrican. Un esquema particular de agregación que se ha sugerido es por *artículos*, *familias* y *tipos*. Los artículos (o unidades de inventario) representan el nivel más fino de detalle y estarían identificados por números distintos de parte. Las familias son grupos de artículos que comparten una preparación común para su manufactura, y los tipos son agrupamientos naturales de familias. Este esquema particular de agregación es bastante general, pero no existe garantía de que funcione en todas las aplicaciones.

En la planeación agregada se supone una *demanda determinista.* Los datos requeridos son pronósticos de demanda durante un horizonte especificado de planeación. Este supuesto no se hace para ser realistas, sino para permitir que el análisis se concentre en los cambios de demanda que sean sistemáticos, más que aleatorios. El objetivo del análisis reside en determinar la cantidad de trabajadores que deben emplearse cada periodo, así como la cantidad de unidades agregadas que deben producirse cada periodo.

El objetivo es minimizar los costos de producción, de nómina, de mantener inventarios y de cambiar el tamaño de la fuerza de trabajo. Los costos de cambiar se llaman generalmente *costos de suavizamiento.* La mayoría de los modelos de planeación agregada descritos en este capítulo suponen que todos los costos son *funciones lineales.* Esto significa que el costo de contratar a un trabajador adicional es igual que el de contratar al anterior, y que el costo de mantener una unidad adicional de inventario es igual al costo de mantener la unidad previa de inventario. Es probable que este supuesto sea una aproximación razonable para la mayor parte de los sistemas reales. No es muy probable que el principal problema en la aplicación de la

metodología de planeación agregada a un caso real pueda ser que la forma de las funciones de costo sea incorrecta; es más probable que la principal dificultad sea estimar bien los costos y demás información requerida.

Se pueden determinar soluciones aproximadas a los problemas de planeación agregada con métodos gráficos. Un programa de producción satisface la demanda si la gráfica de la producción acumulada queda bajo la de la demanda acumulada. Mientras todos los costos sean lineales, se pueden hallar soluciones óptimas[4] con *programación lineal.* Como las soluciones de los problemas de programación lineal pueden ser fracciones, algunas o todas las variables se deberán redondear a enteros. El mejor método para redondear no siempre es obvio. Una vez determinada una solución óptima, debe considerarse su adecuación en el contexto del problema de planeación, ya que hay costos y restricciones intangibles que pueden ser difíciles o imposibles de incluir en forma directa al formular el problema.

También se puede emplear la programación lineal para determinar los niveles óptimos de producción a nivel de artículo individual, pero el tamaño del problema resultante y los requisitos de datos podrían hacer que el método no fuera realista para las empresas grandes. Sin embargo, se pueden usar las formulaciones detalladas de programación lineal a nivel familiar, más que a nivel de artículo individual en las empresas grandes.

En los primeros trabajos de Holt, Modigliani, Muth y Simon (1960) se requería que las funciones de costo fueran cuadráticas. Cuando los costos son cuadráticos, la forma de la política óptima resultante es una función lineal sencilla llamada *regla lineal de decisión.* Sin embargo, el supuesto de funciones de costo cuadráticas probablemente sea demasiado restrictivo para ser preciso en la mayoría de los problemas reales.

El modelo de *coeficientes gerenciales* de Bowman ajusta un modelo matemático razonable del sistema a las decisiones reales tomadas por la gerencia durante un intervalo. Este método tiene la ventaja de incorporar la intuición de la gerencia en el modelo, y motiva al personal gerencial a ser consistente en la futura toma de decisiones.

Los planes agregados serán de poca utilidad para la empresa si no pueden coordinarse con programas detallados por artículo (es decir, el programa maestro de producción). El problema de *desagregar los planes agregados* es difícil, pero se debe atacar si se desea que los planes agregados tengan algún valor para la compañía. Existen algunas formulaciones de programación matemática del problema de desagregación que se han sugerido en las publicaciones, pero todavía está por verse que esos esquemas de desagregación se prueben en la práctica.

Problemas adicionales sobre planeación agregada

29. Se quiere implementar un modelo de planeación agregada para las aplicaciones siguientes. Sugiera una medida agregada de la producción y comente las dificultades que pueden surgir al aplicar la planeación agregada en cada caso.

    a) Planeación de la cantidad de personal docente en una universidad.
    b) Determinación de los requisitos de fuerza de trabajo para una agencia de viajes.
    c) Planeación de los niveles de trabajo y de producción en una planta enlatadora de pescado que procesa sardina, anchoas, arenque y ostras ahumadas.

30. Un gerente de planta de una empresa productora de artículos para cocina a nivel nacional planea niveles de fuerza de trabajo para los tres meses siguientes. Una unidad agregada de

[4] Es decir, óptimas sujetas a errores de redondeo.

producción requiere un promedio de cuatro horas de trabajo. El pronóstico de las demandas para el horizonte de tres meses es el siguiente:

Mes	**Días de trabajo**	**Demanda pronosticada (en unidades agregadas)**
Julio	23	2 400
Agosto	16	3 000
Septiembre	20	800

Suponga que un día normal de trabajo tiene ocho horas. Los costos de contratación son 500 dólares por trabajador, y los de despido son 1 000 dólares por trabajador. Los costos por mantener inventarios son 40 dólares por unidad agregada mantenida por mes. No se permiten faltantes. Suponga que el inventario final al terminar junio era 600, y que el gerente desea tener al menos 800 unidades disponibles al final de septiembre. El nivel de la fuerza de trabajo en junio es de 30 trabajadores.

a) Determine el plan con fuerza de trabajo mínima y constante para los tres meses, y los costos totales de contratación, despido y de mantener inventarios derivados de ese plan.

b) Trace una gráfica de la demanda neta acumulada y úsela para sugerir un plan más eficiente que presente un solo cambio en los niveles de fuerza de trabajo dentro del horizonte de planeación. Calcule el costo del plan y compárelo con el plan propuesto en el inciso *a*). Suponiendo que el segundo plan es más rentable, comente por qué.

31. Una empresa local fabrica juguetes. La demanda proyectada durante los cuatro meses siguientes para un robot de juguete es:

Mes	**Días de trabajo**	**Demanda pronosticada (en unidades agregadas)**
Julio	23	3 825
Agosto	16	7 245
Septiembre	20	2 770
Octubre	22	4 440

Suponga que un día normal de trabajo es de ocho horas. Los costos de contratación son $350 por trabajador y los de despido (incluyendo la indemnización) son $850, también por trabajador. Los costos de mantener el inventario son $4 por unidad agregada almacenada por mes. Suponga que se requiere, en promedio, 1 hora y 40 minutos para que un trabajador arme un juguete. No se permiten faltantes de inventario. Suponga que el inventario final para junio era 600 robots y que el gerente desea tener 800 unidades a la mano, como mínimo, al final de octubre. Suponga que la fuerza de trabajo actual es 35 trabajadores.

a) Encuentre el plan de fuerza de trabajo mínima constante para los cuatro meses, y los costos totales de contratación, despido e inventario de ese plan.

b) Determine el plan que cambie el nivel de fuerza de trabajo cada mes, haciendo coincidir exactamente la demanda y el costo de ese plan.

32. Paris Paint Company planea los niveles de fuerza laboral y de producción para los cuatro trimestres siguientes. El departamento de ventas ha proporcionado al de producción los siguientes pronósticos de la demanda de pinturas durante el año próximo:

Trimestre	Pronóstico de la demanda (en miles de galones)
1	380
2	630
3	220
4	160

Suponga que en la actualidad hay 280 obreros en la empresa. Los trabajadores se contratan durante un trimestre completo como mínimo. Los costos de contratación son $1 200 por trabajador, y los de despido son $2 500 por trabajador. Los costos de inventario son $1 por galón por trimestre. Se estima que un trabajador produce 1 000 galones de pintura cada trimestre.

Suponga que Paris tiene en la actualidad 80 000 galones de pintura en su inventario, y que desea terminar el año con un inventario mínimo de 20 000 galones.

a) Determine el plan de fuerza de trabajo mínima constante y el costo de ese plan. Suponga que no se permiten faltantes.

b) Determine el plan de cero inventarios que contrate y despida trabajadores cada trimestre para hacer coincidir el inventario con la demanda tanto como sea posible, y el costo de ese plan.

c) Si Paris pudiera surtir con retraso el exceso de demanda a un costo de $2 por galón por trimestre, determine un plan con fuerza de trabajo mínima constante que mantenga menos inventario que el que determinó en el inciso *a*), pero que incurra en faltantes en el trimestre 2. Determine el costo del plan nuevo.

33. Considere el problema de Paris Paint descrito en el problema 32. Suponga que la planta tiene capacidad para emplear un máximo de 370 trabajadores. Suponga que los costos por empleado trabajando en tiempo normal son $12.50 por hora. Suponga días de trabajo de siete horas, semanas de cinco días y meses de cuatro semanas. El tiempo extra se paga 50% más. Se puede conseguir subcontratación a un costo de $7 por galón de pintura producido. El tiempo extra se limita a tres horas por empleado por día, y no se pueden subcontratar más de 100 000 galones en trimestre alguno. Proponga una política que cumpla con la demanda y el costo de esa política. Suponga que no se permiten faltantes de inventario.

34. *a*) Formule el problema 32 como un programa lineal. Suponga que no se permiten faltantes de inventario.

b) Resuelva el programa lineal. Redondee las variables y determine el costo del plan resultante.

35. *a*) Formule el problema 33 como un programa lineal.

b) Resuelva el programa lineal. Redondee las variables y determine el costo del plan resultante.



36. Mr. Meadows Cookie Company puede contar con pronósticos precisos para 12 meses, con base en sus pedidos en firme. Esos pronósticos y la cantidad de días de trabajo por mes son:

Mes	Pronóstico de la demanda (en miles de galletas)	Días de trabajo
1	850	26
2	1 260	24
3	510	20
4	980	18
5	770	22
6	850	23
7	1 050	14
8	1 550	21
9	1 350	23
10	1 000	24
11	970	21
12	680	13

Durante un periodo de 46 días, y con una plantilla de 120 trabajadores, la empresa produjo 1 700 000 galletas. Suponga que hay 100 trabajadores empleados al principio del mes 1 y un inventario inicial igual a cero.

a) Calcule la fuerza de trabajo constante mínima para cumplir con la demanda mensual.

b) Suponga que $c_1 = .10$ dólares por galleta por mes, $c_H = \$100$ dólares y $c_F = \$200$. Incluya columnas que expresen el inventario acumulado disponible y el costo de inventario. ¿Cuál es el costo del plan de fuerza de trabajo constante?

c) Elabore un plan que cambie el nivel de la fuerza de trabajo para cumplir con la demanda en la forma más exacta posible. Al establecer la lógica de sus cálculos, asegúrese de que el inventario no pase a valores negativos en ningún mes. Calcule el costo de este plan.

37. La compañía Yeasty Brewing produce una cerveza de consumo local, de marca Iron Stomach. Las ventas de cerveza tienen algo de estacionalidad, y Yeasty debe establecer el plan de sus niveles de producción y fuerza de trabajo el 31 de marzo, para los seis meses siguientes. Los pronósticos de demanda son:

Mes	Días hábiles	Demanda pronosticada (en cientos de cartones)
Abril	11	85
Mayo	22	93
Junio	20	122
Julio	23	176
Agosto	16	140
Septiembre	20	63

El 31 de marzo, Yeasty tenía 86 trabajadores en su nómina. Durante un periodo de 26 días de trabajo, cuando hay 100 trabajadores en nómina, Yeasty puede producir 12 000 cartones de cerveza. El costo de contratar cada trabajador es $125, y el de despido es $300 por trabajador. Mantener el inventario cuesta $0.75 por cartón por mes.

El 31 de marzo, Yeasty espera tener 4 500 cartones en almacén, y desea mantener un inventario mínimo de reserva de 1 000 cartones al final de cada mes. Desea comenzar octubre con 3 000 cartones disponibles.

a) Con base en esta información, elabore el plan de fuerza de trabajo constante mínima para Yeasty durante los seis meses, y calcule los costos de contratación, despido e inventario asociados con ese plan.

b) Suponga que se necesita un mes para adiestrar a un trabajador nuevo. ¿Cómo afectará esto a su solución?

c) Suponga que 10 es la cantidad máxima de trabajadores que la empresa puede contratar en un mes. ¿Cómo afectará eso a su solución en el inciso *a*)?

d) Determine el plan de cero inventarios y su costo [puede hacer caso omiso de las condiciones en los incisos *b*) y *c*)].

38. *a*) Formule el problema de planear los niveles de producción de Yeasty, que se describió en el problema 37, como un programa lineal. Puede hacer caso omiso de las condiciones en los incisos *b*) y *c*).

b) Resuelva el programa lineal resultante. Redondee las variables adecuadas y determine el costo de su solución.

c) Suponga que Yeasty no desea despedir a ningún trabajador. ¿Cuál es el plan óptimo sujeto a esta restricción?



39. Una empacadora local vende verduras enlatadas a una cadena de supermercados. Una caja normal de verduras enlatadas requiere un promedio de .2 días de mano de obra para ser producida. El inventario agregado disponible al final de junio es 800 cajas. Se puede predecir con exactitud la demanda de las cajas durante unos 18 meses, con base en los pedidos que recibe la empresa. Las demandas pronosticadas para los 18 meses siguientes son:

Mes	**Demanda pronosticada (cientos de cajas)**	**Días de trabajo**	**Mes**	**Demanda pronosticada (cientos de cajas)**	**Días de trabajo**
Julio	23	21	Abril	29	20
Agosto	28	14	Mayo	33	22
Septiembre	42	20	Junio	31	21
Octubre	26	23	Julio	20	18
Noviembre	29	18	Agosto	16	14
Diciembre	58	10	Septiembre	33	20
Enero	19	20	Octubre	35	23
Febrero	17	14	Noviembre	28	18
Marzo	25	20	Diciembre	50	10

En la actualidad, la empresa tiene 25 trabajadores. Contratar y entrenar a un trabajador cuesta $1 000 dólares, y el costo de despedirlo es de $1 500. La empacadora estima sus costos en $2.80 por almacenamiento de una caja durante un mes. La empresa desea tener 1 500 cajas en inventario al final del horizonte de planeación de 18 meses.

a) Haga una hoja de trabajo para determinar el plan agregado de fuerza de trabajo mínima constante y determine el costo total de ese plan.

b) Elabore una hoja de trabajo para determinar un plan donde se contraten y despidan trabajadores mensualmente para minimizar los costos de inventario. Calcule también el costo de ese plan.

c) Grafique la demanda neta acumulada y, con la gráfica, proponga un plan que cambie el nivel de fuerza de trabajo exactamente una vez durante los 18 meses, y que tenga un costo menor que cualquiera de los planes considerados en los incisos *a*) y *b*).



40. En este problema se aplicará la regla lineal de producción de Holt, Modigliani, Muth y Simon que apareció en el problema 24. Diseñe una hoja de cálculo para determinar P_t y W_t a partir de los pronósticos de demanda. En particular, suponga que al final de diciembre de 1995 la fuerza de trabajo era de 180, y que el inventario era de 45 unidades agregadas.

a) Con los pronósticos de la demanda para los meses enero a diciembre de 1996 que se mencionan en el problema 24, determine los niveles recomendados de producción agregada y el tamaño de fuerza de trabajo para enero.

**b*) Diseñe su hoja de cálculo de tal manera que pueda actualizar con eficiencia los niveles recomendados cuando se realizan las demandas. En particular, suponga que la demanda para enero fue realmente 175 y que el pronóstico para enero de 1997 fue 130. ¿Cuáles son los niveles recomendados de producción agregada y de tamaño de fuerza de trabajo para febrero? (Sugerencia: Almacene los pronósticos y coeficientes de demanda en columnas separadas y modifique la fórmula de cálculo a medida que los datos nuevos se agreguen a los anteriores.)

**c*) Las realizaciones de la demanda y los nuevos pronósticos de la demanda para el resto de 1996 son los de la tabla siguiente. Calcule los valores actualizados de P_t y W_t cada mes.

Mes	Demanda real (1996)	Nuevo pronóstico (1997)
Febrero	148	160
Marzo	200	175
Abril	190	120
Mayo	180	190
Junio	130	230
Julio	95	280
Agosto	120	210
Septiembre	68	110
Octubre	88	95
Noviembre	75	92
Diciembre	60	77

Apéndice **3-A**

Glosario de notación para el capítulo 3

α = Constante de suavizamiento para la producción y la demanda, usada en el modelo de Bowman.

β = Constante de suavizamiento para el inventario, usada en el modelo de Bowman.

c_F = Costo de despedir un trabajador.

c_H = Costo de contratar un trabajador.

c_I = Costo de mantener una unidad de inventario durante un periodo.

c_O = Costo de producir una unidad en tiempo extra.

c_P = Costo de penalización por pedidos devueltos.

c_R = Costo de producir una unidad en tiempo normal.

c_S = Costo de subcontratar una unidad de producción.

c_U = Costo de tiempo inactivo por unidad de producción.

D_t = Demanda pronosticada para el periodo t.

F_t = Cantidad de trabajadores despedidos en el periodo t.

H_t = Cantidad de trabajadores contratados en el periodo t.

I_t = Nivel de inventario en el periodo t.

K = Cantidad de unidades agregadas producidas por un trabajador en un día.

λ_j = Demanda anual para la familia j (vea la sección 3.9).

n_t = Número de días de producción en el periodo t.

O_t = Producción durante tiempo extra, en unidades.

P_t = Nivel de producción en el periodo t.

S_t = Número de unidades subcontratadas en el exterior.

T = Número de periodos en el horizonte de planeación.

U_t = Tiempo ocioso de trabajadores en unidades de producción.

W_t = Nivel de la fuerza de trabajo en el periodo t.

Bibliografía

Allen, S. J. y E. W. Schuster. "Practical Production Scheduling with Capacity Constraints and Dynamic Demand: Family Planning and Disaggregation". *Production and Inventory Management Journal* 35, 4(1994), pp. 15-20.

Axsater, S. "Aggregation of Product Data for Hierarchical Production Planning". *Operations Research* 29(1981), pp. 744-756.

Bergstrom, G. L. y B. E. Smith. "Multi-item Production Planning—An Extension of the HMMS Rules". *Management Science* 16(1970), pp. 614-629.

Bitran, G. R. y A. Hax. "Disaggregation and Resource Allocation Using Convex Knapsack Problems with Bounded Variables". *Management Science* 27(1981), pp. 431-441.

Bowman, E. H. "Production Scheduling by the Transportation Method of Linear Programming". *Operations Research* 4(1956), pp. 100-103.

Bowman, E. H. "Consistency and Optimality in Managerial Decision Making". *Management Science* 9(1963), pp. 310-321.

Buffa, E. S. *Modern Production Operations Management*, septima edición. Nueva York: John Wiley & Sons, 1983.

Buffa, E. S. y W. H. Taubert. *Production-Inventory Systems: Planning and Control*, edición revisada, Homewood, IL: Richard D. Irwin, 1972.

"Caterpillar Tractor". Harvard Business School Case 385-276(1985).

Chung, C. y L. Krajewski. "Planning Horizons for Master Production Scheduling". *Journal of Operations Management*, agosto de 1981, pp. 389-406.

Cohen, M. A., M. L. Fisher y J. Jaikumar. "International Manufacturing and Distribution Networks". En *Managing International Manufacturing*, K. Ferdows, editor, pp. 67-93. Amsterdam: Norte de Holanda, 1989.

Damon, W. W. y R. Schramm. "A Simultaneous Decision Model for Production, Marketing, and Finance". *Management Science* 19(1972), pp. 161-172.

Ebert, R. J. "Aggregate Planning with Learning Curve Productivity". *Management Science* 23(1976), pp. 171-182.

Erenguc, S. y S. Tufekci. "A Transportation Type Aggregate Production Model with Bounds on Inventory and Backordering". *European Journal of Operations Research* 35(1988), pp. 414-425.

Ferdows, K., editor, *Managing International Manufacturing.* Amsterdam: Norte de Holanda, 1989.

Hansmann, F. y S. W. Hess. "A Linear Programming Approach to Production and Empoloyment Scheduling". *Management Technology* 1, 1960, pp. 46-51.

Hax, A. C. y D. Candea, *Production and Inventory Management.* Englewood Cliffs, NJ: Prentice Hall, 1984.

Hax, A. C. y H. C. Meal. "Hierarchical Integration of Production Planning and Scheduling". En *TIMS Studies in Management Science.* Volumen 1. *Logistics*, M. Geisler, editor, Nueva York: Elsevier, 1975.

Hillier, F. S. y G. J. Lieberman. *Introduction to Operations Research*, quinta edición, San Francisco: Holden Day, 1990.

Holt, C. C., F. Modigliani y J. F. Muth. "Derivation of a Linear Decision Rule for Production and Employment". *Management Science* 2(1956), pp. 159-177.

Holt, C. C., F. Modigliani, J. F. Muth y H. A. Simon. *Planning Production, Inventories, and Workforce.* Englewood Cliffs. NJ: Prentice Hall, 1960.

Holt, C. C., F. Modigliani y H. A. Simon. "A Linear Decision Rule for Employment and Production Scheduling". *Management Science* 2(1955), pp. 1-30.

Huchzermeier, A. y M. A. Cohen. "Valuing Operational Flexibility under Exchange Rate Risk". *Operations Research* 44(1996), pp. 100-113.

Kamien, M. I. y L. Li. "Subcontracting. Coordination, Flexibility, and Production Smoothing in Aggregate Planning". *Management Science* 36(1990), pp. 1352-1363.

Kogut, B. y N. Kulatilaka. "Operating Flexibility, Global Manufacturing, and the Option Value of a Multinational Network". *Management Science* 40(1994), pp. 123-139.

Leitch, R. A. "Marketing Strategy and Optimal Production Schedule". *Management Science* 20(1974), pp. 903-911.

McGrath, M. E. y R. B. Bequillard. "International Manufacturing Strategies and Infrastructural Considerations in the Electronics Industry". En *Managing International Manufacturing*, K. Ferdows, editor, pp. 23-40. Amsterdam: Norte de Holanda, 1989.

Newson, E. F. P. "Multi-Item Lot Size Scheduling by Heuristic, Part 1: With Fixed Resources". *Management Science* 21(1975a), pp. 1186-1193.

Newson, E. F. P. "Multi-Item Lot Size Scheduling by Heuristic, Part 2: With Fixed Resources". *Management Science* 21(1975b), pp. 1194-1205.

Schrage, L. *Linear, Integer, and Quadratic Programming with LINDO.* Palo Alto, CA: Scientific Press, 1984.

Schwarz, L. B. y R. E. Johnson. "An Appraisal of the Empirical Performance of the Linear Decision Rule for Aggregate Planning". *Mannagement Science* 24(1978), pp. 844-849.

Silver, E. A. y R. Peterson. *Decision Systems for Inventory Management and Production Planning.* Segunda edición, Nueva York: John Wiley & Sons, 1985.

Taubert, W. H. "A Search Decision Rule for the Aggregate Scheduling Problem". *Management Science* 14(1968), pp. B343-359.

Vollman, T. E., W. L. Berry y D. C. Whybark. *Manufacturing, Planning, and Control Systems.* Tercera edición, Nueva York: McGraw-Hill/Irwin, 1992.

Zoller, K. "Optimal Disaggregation of Aggregate Production Plans". *Management Science* 17(1971), pp. B533-549.

Suplemento **Uno**

Programación lineal

S1.1 INTRODUCCIÓN

La programación lineal es una técnica matemática para resolver una clase amplia de problemas de optimización. En esos problemas se requiere maximizar o minimizar una función lineal de n variables reales, sujeta a m restricciones. Se pueden formular y resolver una gran cantidad de problemas reales con la programación lineal. Entre ellos se incluyen:

1. Programación de personal.
2. Algunas variedades de problemas de formulación de mezclas, por ejemplo, de alimento para ganado, condimentos alimenticios, helados y fabricación de acero.
3. Control de inventarios y planeación de producción.
4. Problemas de distribución y logística.
5. Problemas de asignación.

Hoy en las computadoras se resuelven fácilmente problemas con miles de variables y restricciones. La programación lineal fue desarrollada para resolver problemas logísticos durante la Segunda Guerra Mundial. George Dantzig, matemático de la RAND Corporation en esa época, desarrolló un procedimiento de solución al que llamó método simplex. El mismo Dantzig se sorprendió porque el método resultó ser muy eficiente para resolver con rapidez problemas grandes. Este hecho, junto con el desarrollo simultáneo del cómputo electrónico, ubicó la programación lineal como una herramienta importante en la administración de logística. El éxito de la programación lineal en la industria impulsó el desarrollo de las disciplinas de investigación de operaciones y la ciencia administrativa. El método simplex ha pasado la prueba del tiempo. Sólo hasta fechas recientes se desarrolla otro método que podría tener potencialmente más eficacia que el simplex para resolver problemas lineales muy grandes y con estructuras especiales. Dicho método, llamado algoritmo de Karmarkar, lleva el apellido del matemático que lo concibió en los Laboratorios Bell.

En la sección S1.2 consideramos un problema característico de manufactura, que formularemos y resolveremos usando programación lineal. Después investigaremos cómo se resuelven problemas pequeños (que tienen exactamente dos variables de decisión) en forma gráfica, y cómo se resuelven problemas grandes usando computadora.

S1.2 UN PROBLEMA PROTOTIPO DE PROGRAMACIÓN LINEAL

Ejemplo *S1.1* Sidneyville fabrica muebles de oficina y para el hogar. La División Oficina produce dos escritorios, el de tapa corrediza o de cierre y el normal. Los fabrica en su planta en las afueras de Medford, Oregon, usando una selección de maderas. Éstas se cortan a un espesor uniforme de 1 pulgada. Por esta razón, la madera se mide en metros cuadrados. Un escritorio de cierre requiere 10 metros cuadrados de pino, 4 de cedro y 15 de arce. Para un escritorio normal se requieren 20 metros cuadrados de pino, 15 de cedro

y 10 de arce. Los escritorios producen ganancias respectivas de 115 dólares y 90 dólares por venta. En la actualidad, la empresa dispone de 200 metros cuadrados de pino, 128 de cedro y 220 de arce. Han recabado pedidos para ambos escritorios y les gustaría producir una cantidad de piezas con cierre y normales que maximice su ganancia. ¿Cuántos escritorios deben producir de cada uno?

Solución El primer paso para formular un problema como programa lineal es identificar las variables de decisión. En este caso se requieren dos decisiones: la cantidad de escritorios de cierre y la de escritorios normales. Debemos asignar símbolos a cada una de esas variables de decisión.

Sean

x_1 = Número de escritorios de cierre que deben producirse.

x_2 = Número de escritorios normales que deben producirse.

Ahora que hemos identificado las variables de decisión, el siguiente paso es identificar la función objetivo y las restricciones. La función objetivo es la cantidad que se desea maximizar o minimizar. El objetivo es maximizar las ganancias, de modo que la función objetivo es igual a la ganancia total cuando se producen x_1 escritorios de cierre y x_2 normales. Cada escritorio de cierre contribuye con $115 a la ganancia, de modo que la contribución total a la ganancia debida a todos los escritorios de cierre que se produzcan es $115x_1$. De igual manera, la contribución a la ganancia debida a todos los escritorios normales es $90x_2$. Por lo tanto, la ganancia total es $115x_1 + 90x_2$. Es a esto a lo que se le llama función objetivo.

El siguiente paso es identificar las restricciones. La cantidad de escritorios que puede producir Sidneyville está limitada por la cantidad de madera disponible. Los tres tipos de madera constituyen los recursos críticos. Para obtener las restricciones necesitamos deducir expresiones para calcular la cantidad de cada tipo de madera que se consume al construir x_1 escritorios de cierre y x_2 escritorios normales. Por consiguiente, esas expresiones están limitadas por la cantidad de cada tipo de madera disponible.

La cantidad de metros cuadrados de pino que se usa para fabricar x_1 escritorios de cierre es $10x_1$. La cantidad de metros cuadrados de pino usada para fabricar x_2 escritorios normales es $20x_2$. Entonces, la cantidad total de pino consumido, en metros cuadrados, es $10x_1 + 20x_2$. Esta cantidad no puede ser mayor a la de metros cuadrados de pino disponibles, que es 200. Así obtenemos la primera restricción:

$$10x_1 + 20x_2 \leq 200.$$

Las otras dos restricciones son parecidas. La segunda es para asegurar que no se rebase la disponibilidad de cedro. Cada escritorio de cierre requiere 4 metros cuadrados de cedro, y x_1 escritorios de cierre requieren $4x_1$ metros cuadrados de cedro. Cada escritorio normal requiere 16 metros cuadrados de cedro, así que x_2 escritorios normales requieren $16x_2$ metros cuadrados de cedro. Por lo anterior, la restricción de la disponibilidad de cedro es

$$4x_1 + 16x_2 \leq 128.$$

De la misma forma, la restricción final para asegurar que no se rebase el arce disponible es

$$15x_1 + 10x_2 \leq 220.$$

Como no se pueden producir cantidades negativas de escritorios, también se incluyen las restricciones de no negatividad:

$$x_1 \geq 0,$$
$$x_2 \geq 0.$$

Ya hemos construido la formulación completa de programación lineal para el problema de Sidneyville. De lo que se trata es de calcular los valores de x_1 y x_2 que maximicen $115x_1 + 90x_2$ sujeta a las restricciones.

$$\begin{aligned} 10x_1 + 20x_2 &\leq 200, \\ 4x_1 + 16x_2 &\leq 128, \\ 15x_1 + 10x_2 &\leq 220, \\ x_1, x_2 &\geq 0. \end{aligned}$$

TABLA S1-1 **Salida parcial de computadora para el problema de Sidneyville**

```
LP OPTIMUM FOUND AT STEP 2

OBJECTIVE FUNCTION VALUE

1) 1740.000000

VARIABLE          VALUE         REDUCED COST
   X1           12.000000         .000000
   X2            4.000000         .000000
```

A continuación consideraremos cómo resolver este tipo de problemas. Mencionaremos en forma somera la teoría en la que se basa la técnica de solución llamada método simplex. Sin embargo, como los problemas de programación lineal ya casi nunca se resuelven manualmente, el lector no necesitará la mecánica de ese método para aplicarlo. Deberá conocer cómo formular los problemas en forma de programas lineales, capturar las formulaciones en la computadora, reconocer los problemas especiales y analizar los resultados de cómputo.

En la tabla S1-1 se presenta una salida típica de computadora.

Este resultado indica que en la solución óptima, el valor x_1 es 12 y el valor x_2 es 14. Esto es, Sidneyville debe producir exactamente 12 escritorios de cierre y 4 escritorios normales. El valor de la función objetivo en la solución óptima es $1 740 00. Ésta es la solución óptima (única), lo que significa que cualquier otro par de valores de x_1 y x_2 darán como resultado una ganancia menor, que no se cumplan las restricciones, o ambas cosas.

El gerente de planeación de la producción de Sidneyville no confía en las matemáticas. Cuando se le presentó esta solución, su respuesta fue "sólo hay un problema con este plan de producción. Tenemos un especialista que hace el cierre de los escritorios que tienen dicho sistema. Sólo puede hacer cuatro escritorios diarios, y deseamos estar listos para embarcar en dos días. No hay forma de producir 12 de esos escritorios en dos días. "¡Ya sabía que esas matemáticas son fantasía!"

El gerente estaba equivocado. El problema no es que la formulación esté incorrecta, sino que no incluye todas las restricciones relevantes, puesto que ahora se sabe que las horas de trabajo eran un recurso crítico. La lección en este caso es que, para que la solución final tenga sentido, el modelo debe incluir *todas* las restricciones relevantes.

S1.3 FORMULACIÓN DEL PROBLEMA GENERAL

El problema de Sidneyville es un ejemplo de programación lineal en el que hay dos variables de decisión y tres restricciones. Estos problemas pueden tener cualquier cantidad de variables de decisión y de restricciones. Supóngase que hay n variables de decisión, representadas por x_1, $x_2, \ldots, x_n$, sujetas a m restricciones de recursos. Se puede escribir el problema de maximizar la función objetivo sujeta a las restricciones de la siguiente forma:

$$\text{Maximizar } c_1x_1 + c_2x_2 + \cdots + c_nx_n,$$

$$\begin{aligned} \text{sujeta a } \; a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\le b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &\le b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\le b_m, \\ x_1, x_2, \ldots, x_n &\ge 0. \end{aligned}$$

Las constantes $c_1, c_2, \ldots, c_n$ se interpretan como los coeficientes de ganancia por unidad de producción de las actividades $x_1, x_2, \ldots, x_n$; a_{ij} es la cantidad del recurso i consumida por una unidad de actividad j; y b_1 es la cantidad disponible del recurso i, para $i = 1, \ldots, n$. Se

requiere que las constantes $b_1, \ldots, b_m$ no sean negativas. Esta formulación particular se aplica a problemas en los que se desea maximizar la ganancia, con apego a las restricciones de los recursos de las variables. Sin embargo, la programación lineal se puede usar para resolver una variedad mucho más amplia de problemas. Más adelante se describen otras formulaciones posibles.

Definición de términos de uso común

1. *Función objetivo*. Es la cantidad que se desea maximizar o minimizar. En la formulación anterior, la función objetivo es la expresión $c_1x_1 + c_2x_2 + \cdots + c_nx_n$. En aplicaciones en las empresas se suele minimizar el costo o maximizar la ganancia. Usaremos las abreviaturas "min" para un problema de minimización y "max" para uno de maximización.

2. *Restricciones.* Cada restricción es una desigualdad o una ecuación lineal; esto es, una combinación lineal de las variables del problema seguida de un operador relacional ($\leq$, $=$ o $\geq$) al que sigue una constante no negativa. Aunque la formulación anterior tiene $\leq$ en todas sus restricciones, también son comunes las restricciones de los tipos $\geq$ e $=$. Por ejemplo, suponga que hay un convenio contractual que establece una cantidad mínima diaria de horas de trabajo. Esto daría como resultado una restricción del tipo $\geq$.

3. *Lado derecho.* Constante que sigue al operador relacional en una restricción. Las constantes $b_1, b_2, \ldots, b_m$ son los lados derechos en la formulación anterior. Se requiere que esas constantes sean números no negativos. Sin embargo, no se requiere que las constantes a_{ij} sean no negativas. Esto significa que cualquier restricción puede escribirse con un lado derecho no negativo, multiplicando todo por la constante -1 cuando el lado derecho sea negativo. Veamos el siguiente ejemplo sencillo. Suponga que cuando se formula un problema como programa lineal se obtiene la restricción

$$4x_1 - 2x_2 \leq -5.$$

Como el lado derecho es negativo, esa restricción no es legal. Sin embargo, si multiplicamos todo por -1, esta restricción se transforma en

$$-4x_1 + 2x_2 \geq 5,$$

la cual es aceptable.

4. *Región factible.* Es el conjunto de valores de las variables de decisión, $x_1, x_2, \ldots, x_n$ que satisfacen las restricciones. Como cada una de las restricciones se genera mediante una ecuación o desigualdad lineal, la región factible tiene una estructura particular. El término técnico para indicar esa estructura es *polítope convexa*. En dos dimensiones, una polítope convexa es un conjunto cuyas fronteras son líneas rectas. En tres dimensiones, las fronteras se forman con planos. Un conjunto convexo se caracteriza así: se eligen dos puntos cualesquiera del conjunto y se unen con una recta. Esa recta está totalmente dentro del conjunto.

5. *Puntos extremos.* Debido a la estructura de la región factible, habrá una cantidad finita de puntos factibles cuya propiedad es que no pueden expresarse como una combinación lineal de cualquier otro conjunto de puntos factibles. Esos puntos se llaman puntos extremos, o puntos de esquina, y desempeñan un papel importante en la programación lineal. El concepto de los puntos extremos se aclarará al describir las soluciones gráficas.

6. *Solución factible.* Es un conjunto particular de valores de las variables de decisión que satisface las restricciones. Una solución factible también es un punto en la región factible. Puede ser un punto extremo o un punto interior.

7. *Solución óptima.* Es la solución factible que maximiza o minimiza la función objetivo. En algunos casos, la solución óptima puede no ser única. En esos casos habrá una cantidad infinita de soluciones óptimas.

Propiedades de los programas lineales

La programación lineal es una herramienta muy poderosa. Con ella muchos problemas reales se han formulado y resuelto correctamente. Sin embargo, para usar bien el método deben tenerse en cuenta sus limitaciones. Dos propiedades importantes de los programas lineales son la linealidad y la continuidad. Varios problemas que parecerían poder solucionarse mediante la programación no cumplen con una o ambas de esas propiedades.

Linealidad

Los problemas de optimización se pueden formular como programas lineales sólo cuando *a*) el objetivo se puede expresar como una función lineal de las variables de decisión, y *b*) cuando todas las restricciones se pueden expresar como funciones lineales de las variables de decisión.

La linealidad implica que las cantidades cambien en proporciones fijas. Por ejemplo, si cuesta $10 producir una unidad, debe costar $20 producir dos, y $100 producir 10. Si una onza de jugo de naranja proporciona 30 mg de vitamina C, tres onzas deben proporcionar 90 mg. La linealidad debe ser válida en la función objetivo y en las restricciones. En la función objetivo esto significa que la ganancia o el costo por unidad deben ser los mismos, independientemente de la cantidad de unidades. En las restricciones, linealidad significa que la cantidad de cada recurso consumido es igual por unidad, ya sea que se produzcan muchas unidades o una sola.

Sin embargo, en el mundo real se observan con frecuencia relaciones no lineales: las economías de producción a escala se traducen en que el costo marginal de producir una unidad disminuye a medida que aumenta la cantidad de unidades producidas. Cuando esto sucede, el costo de producción es una función no lineal de la cantidad de unidades producidas. Un ejemplo de economías de escala es un costo fijo de preparación. La fórmula para la cantidad económica de pedido que se describirá en el capítulo 4 indica que el tamaño del lote aumenta en función de la raíz cuadrada de la tasa demandada. Por consiguiente, la cantidad económica de pedido es una función no lineal de la demanda. Cuando la función objetivo o una restricción son funciones no lineales de las variables de decisión, el problema es de programación no lineal, y no se puede resolver con programación lineal.[1]

Continuidad

Significa que las variables de decisión deben ser continuas (esto es, capaces de asumir cualquier valor no negativo), y no tener valores discretos o enteros. Tal restricción puede ser grave. Resolver muchos problemas tiene sentido sólo si las variables de decisión asumen un valor entero. En particular, el ejemplo S1.1 no es, propiamente dicho, un problema de programación lineal, porque el número de escritorios producidos debe ser entero. (Fuimos afortunados porque en este caso la solución óptima tuvo valor entero.) Se puede pensar que la solución entera óptima es igual a la solución continua redondeada al entero más próximo. Desafortunadamente, no siempre sucede así. En primer lugar, el redondeo tal vez origine infactibilidad; esto es, que la solución redondeada puede quedar fuera de la región factible. En segundo lugar, aun cuando la solución redondeada sea factible, quizás no sea la óptima. Puede suceder que la solución entera óptima esté en una parte totalmente aparte de la región factible de la solución que la solución redondeada de programación.

Para que el lector tenga una idea de las dificultades que pueden surgir cuando la solución debe tener valores enteros, veamos el ejemplo S1.1. Suponga que la ganancia por vender escritorios de cierre fuera de $150, y no de $115. Entonces el objetivo sería maximizar $150x_1 + 90x_2$ sujeta al mismo conjunto de restricciones. La solución óptima de programación lineal es

$$x_1 = 14.666666\ldots.$$
$$x_2 = 0.$$

[1] En algunos casos se pueden resolver problemas convexos de programación con programación lineal, aproximando la función objetivo con una función lineal en intervalos. Un ejemplo de este método se ve al final de la sección 3.5.

Al redondear la solución al entero más próximo se obtiene $x_1 = 15$ y $x_2 = 0$, que no es factible. Al sustituir esos valores en la restricción final se obtiene un requerimiento de 225 metros cuadrados de arce. Sin embargo, sólo se dispone de 220 metros cuadrados. Al redondear x_1 a 14, se obtiene una solución factible, pero subóptima. Con $x_1 = 14$ y $x_2 = 0$, quedan disponibles 10 pies de arce, suficientes para fabricar un escritorio normal. En este caso, la solución entera óptima es $x_1 = 14$ y $x_2 = 1$.

Cuando las variables de decisión deben tener valores enteros se dice que el problema es de programación lineal entera, o programación entera. Determinar soluciones enteras óptimas para programas lineales puede ser muy tardado, aun para problemas de tamaño modesto. Sin embargo, el paquete LINDO ofrece una opción para definir algunas, o todas las variables del problema como de valor entero. Este paquete funciona bien en la solución de problemas pequeños de programación lineal. Para problemas mayores se debería usar un programa informático cuyo propósito sea resolver problemas de programación entera. En algunos casos, en especial cuando los valores de las variables son relativamente grandes, el redondeo cuidadoso de la solución de programación lineal debe dar resultados aceptables.

S1.4 SOLUCIÓN GRÁFICA DE PROBLEMAS DE PROGRAMACIÓN LINEAL

Graficación de desigualdades lineales

En esta sección mostraremos cómo resolver problemas de programación lineal con dos variables, en forma gráfica. Aunque la mayoría de los problemas reales tienen más de dos variables, la comprensión del procedimiento para resolver programas con dos variables mejorará la percepción de los conceptos sobre los que se base el método simplex.

El primer paso es graficar las desigualdades lineales representadas por las restricciones. Una desigualdad lineal corresponde a todos los puntos del plano a un lado de una línea recta. Por consiguiente, para graficar las desigualdades lineales se siguen dos pasos:

1. Trazar la recta que representa la frontera o cota de la región correspondiente a la restricción, expresada como ecuación.
2. Determinar cuál lado de la recta corresponde a la desigualdad.

Para ilustrar el método, veamos la primera restricción en el ejemplo S1.1:

$$10x_1 + 20x_2 \leq 200.$$

La frontera de la región representada por esta desigualdad es la recta.

$$10x_1 + 20x_2 = 200.$$

La forma más fácil de graficar una recta es determinar las dos intersecciones con los ejes. Se determinan igualando x_2 a cero y despejando x_1 y después igualando x_1 a cero y despejando x_2. Primero, al igualar x_2 a cero se obtiene

$$10x_1 = 200 \qquad \text{o} \qquad x_1 = 20.$$

De la misma manera, al igualar x_1 a cero y despejar x_2 se obtiene

$$20x_2 = 200 \qquad \text{o} \qquad x_2 = 10.$$

Por consiguiente, la recta $10x_1 + 20x_2 = 200$ debe pasar por los puntos (20, 0) (la intersección con el eje x_1) y (0, 10) (la intersección con el eje x_2). En la figura S1-1 se ve una gráfica de esa recta. Ya graficada la línea que define la frontera del semiespacio, se debe determinar cuál lado de la recta corresponde a la desigualdad. Para hacerlo elegimos cualquier punto que *no* esté en la línea, sustituimos los valores de x_1 y x_2 y vemos si la desigualdad se satisface o

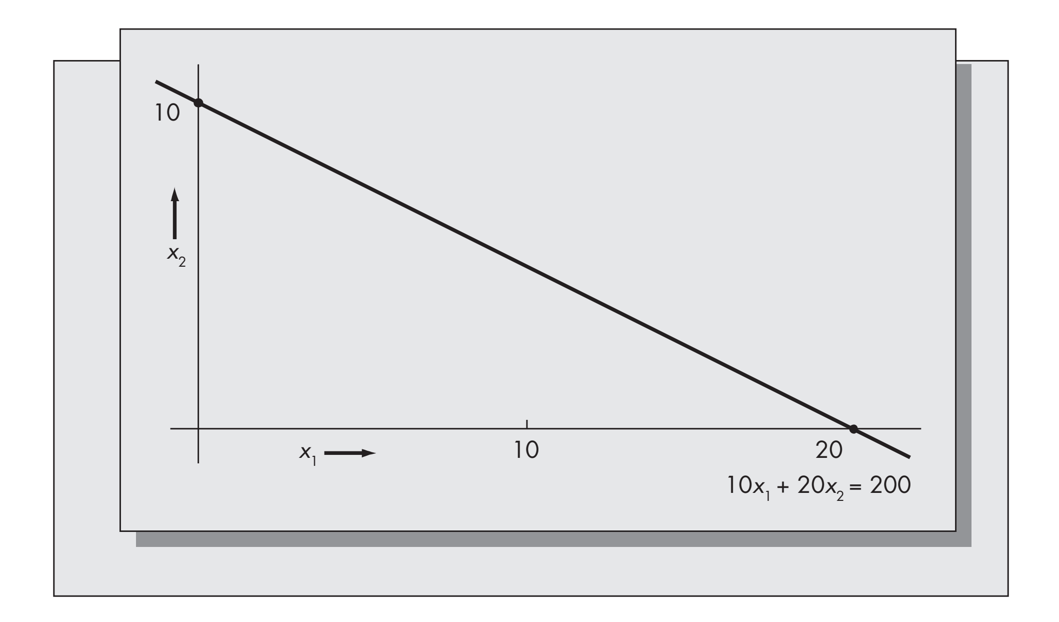


FIGURA S1-1
Graficación de una frontera de restricción.

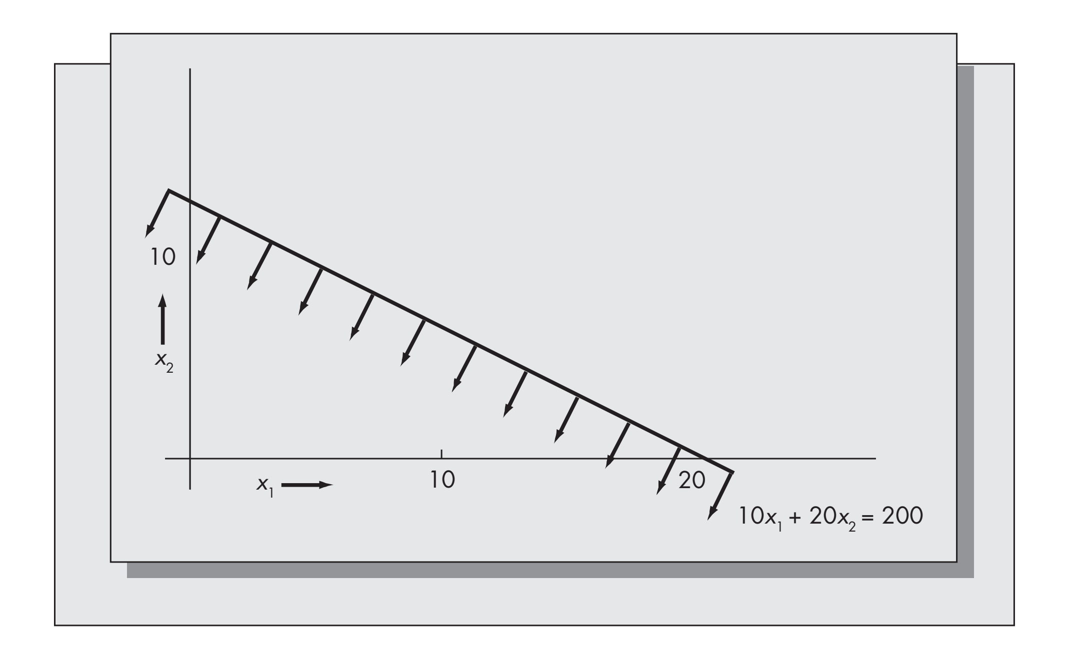


FIGURA S1-2
Semiespacio representando la desigualdad $10x_1 + 20x_2 \leq 200$.

no. Si se satisface, ese punto pertenece al semiespacio; si no, no pertenece al semiespacio. Si la línea de cota no pasa por el origen [es decir, el punto (0, 0)], el método más directo es usar el origen como punto de prueba.

Al sustituir (0, 0) en la desigualdad, se obtiene

$$(10)(0) + (20)(0) = 0 \leq 200.$$

Como 0 es menor que 200, la prueba se satisface y el origen está dentro de la región representada por la desigualdad. Esto quiere decir que la gráfica de la desigualdad incluye a todos los puntos debajo de la línea de frontera en la figura S1-1, como se ve en la figura S1-2.

La prueba con el origen para determinar cuál lado de la recta es el adecuado sólo funciona cuando la misma línea de frontera no pasa por el origen. Si lo hace, habrá que usar algún otro

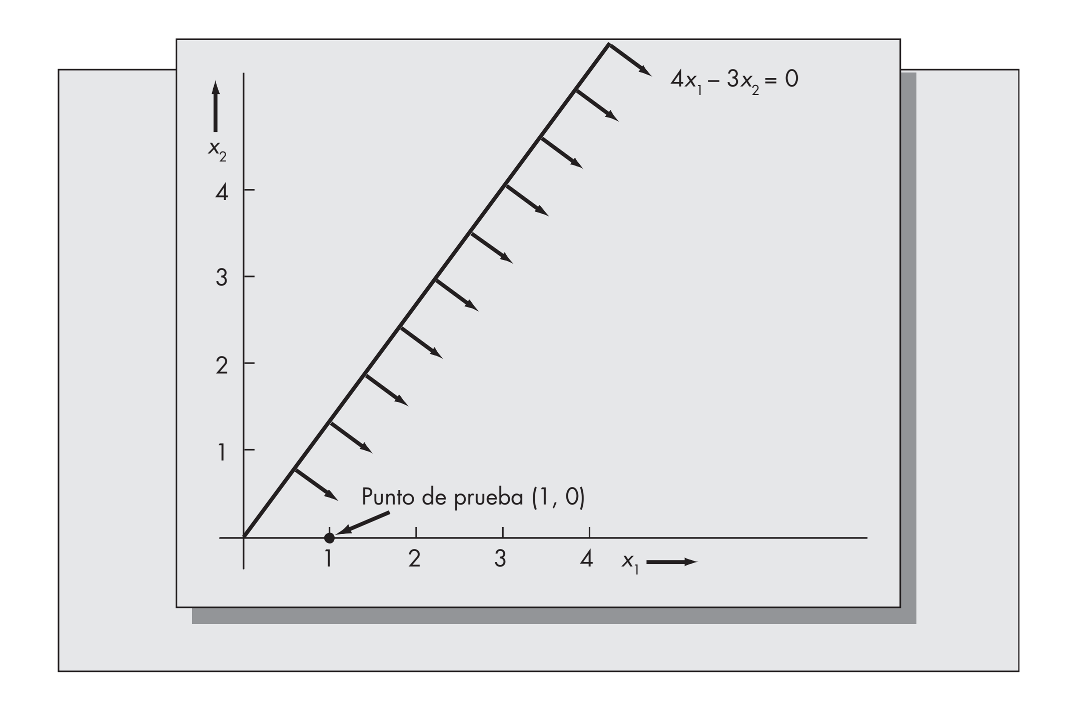


FIGURA S1-3
Graficación de una frontera de restricción que pasa a través del origen.

punto que no esté sobre la línea para realizar la prueba. Por ejemplo, si se tiene la restricción

$$4x_1 - 3x_2 \geq 0.$$

Al tratar de graficar la línea $4x_1 - 3x_2 = 0$, vemos que al sustituir $x_1 = 0$ se obtiene $x_2 = 0$, y si se sustituye $x_2 = 0$ se obtiene $x_1 = 0$. Esto quiere decir que la línea pasa por el origen. Para graficar dicha línea debemos determinar otro punto que esté en ella. Tan sólo se elige cualquier valor de x_1 y se despeja el valor correspondiente de x_2. Por ejemplo, si se sustituye $x_1 = 3$ se obtiene $x_2 = 4$, que significa que el punto (3, 4) está en la recta, al igual que el punto (0, 0). La recta de frontera se grafica uniendo esos puntos, como se aprecia en la figura S1-3.

A continuación determinaremos cuál lado de la desigualdad corresponde a la región de interés. Como se observó anteriormente, la prueba con el origen no funciona cuando la línea pasa por él. Se elige cualquier punto que *no* esté en la línea para la prueba. En este caso, uno que no está en ella es $x_1 = 1$ y $x_2 = 0$. Al sustituir esos valores en la desigualdad se obtiene

$$(4)(1) - 0 = 4 > 0.$$

Se satisface la desigualdad, y así el punto (1, 0) está en la región. La desigualdad corresponde a los puntos debajo de la recta, como se indica en la figura S1-3.

Graficación de la región factible

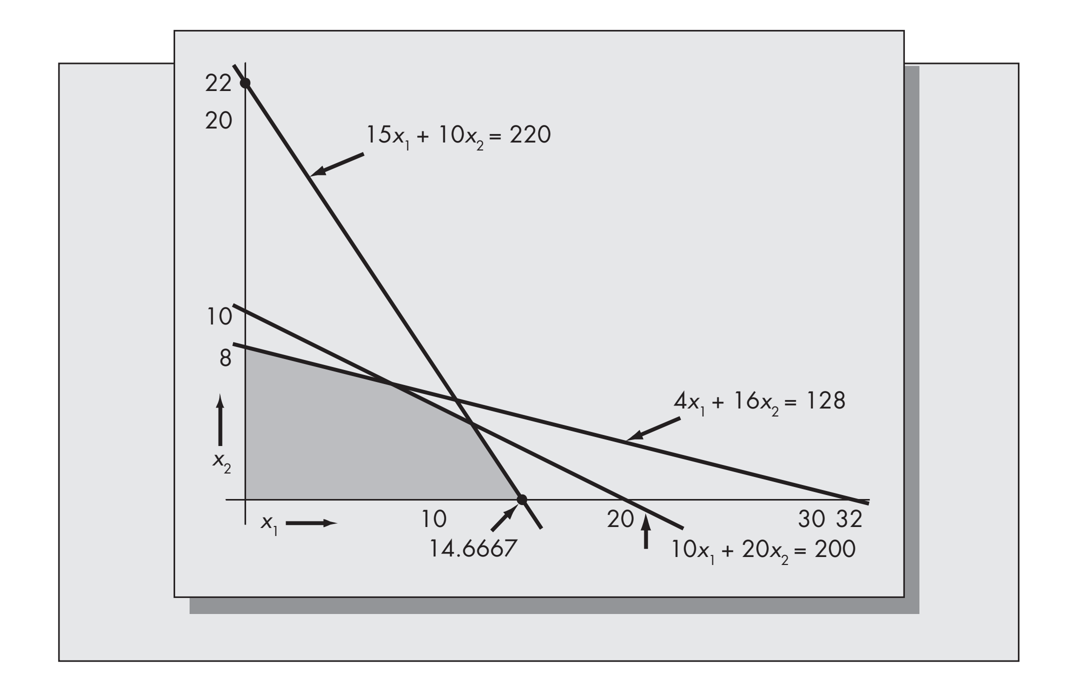
La gráfica de la región factible se determina graficando las desigualdades lineales representadas por las restricciones, y determinando la región de intersección de los semiespacios correspondientes. En esta forma determinaremos una gráfica de la región factible para el problema de Sidneyville.

En la figura S1-2 se ha graficado la región factible correspondiente a la primera restricción. Quedan las otras dos restricciones:

$$4x_1 + 16x_2 \leq 128,$$
$$15x_1 + 10x_2 \leq 220.$$

Los semiespacios que corresponden a esas restricciones se determinan de la misma forma. Primero se grafican las rectas que corresponden a las cotas de región. En el primer caso las

FIGURA S1-4
Región factible para el problema de Sidneyville (Ejemplo S1.1).



intercepciones son $x_1 = 32$ y $x_2 = 8$, y $x_1 = 14.6667$ y $x_2 = 22$ en el segundo caso. Al aplicar la prueba del origen se ve que los semiespacios adecuados son los puntos que se ubican debajo de ambas rectas. Además, también deben incluirse las restricciones de no negatividad, $x_1 \geq 0$ y $x_2 \geq 0$. La región factible que resulta se muestra en la figura S1-4.

Determinación de la solución óptima

La región factible que se aprecia en la figura S1-4 tiene varias propiedades interesantes, que son comunes a todos los problemas de programación lineal. Si se eligen dos soluciones factibles cualesquiera (esto es, puntos en la región) y se unen esos puntos con una recta, la línea que resulta queda por completo dentro de la región, lo que significa que la región factible es un conjunto convexo. Las fronteras de la región son rectas. Esas rectas se intersecan en los llamados puntos extremos. Hay un total de cinco puntos extremos en la región factible representada en la figura S1-4.

Una propiedad importante de los programas lineales es que la solución óptima siempre se presenta en un punto extremo de la región factible. Para el problema de Sidneyville, esto implica que entre la cantidad infinita de soluciones (es decir, puntos) en la región factible, ¡la solución óptima será sólo uno de cinco puntos![2] Esto es cierto, sin importar cuál función objetivo se suponga, lo que significa que se puede determinar la solución óptima para el problema de Sidneyville identificando los cinco puntos extremos, sustituyendo sus coordenadas (x_1, x_2) en la función objetivo y determinando el punto que proporcione la ganancia máxima. Describiremos primero este método, aunque hay un procedimiento gráfico de solución más eficiente.

En la figura S1-4 se observa que uno de los puntos extremos es el origen (0, 0). Otro es la intercepción x_2, que corresponde a la restricción $4x_1 + 16x_2 = 128$, que es (0, 8). Un tercero es la intercepción x_1, que corresponde a la restricción $15x_1 + 10x_2 = 220$, que es (14.6667, 0). Los otros dos puntos extremos corresponden a las intersecciones de pares de líneas de frontera. Se determinan resolviendo simultáneamente las ecuaciones correspondientes a esas fronteras.

[2] Puede suceder que dos puntos extremos sean óptimos, en cuyo caso todos los puntos en la recta que los une también son óptimos.

Primero resolveremos simultáneamente las ecuaciones

$$4x_1 + 16x_2 = 128,$$
$$10x_1 + 20x_2 = 200.$$

Esas ecuaciones pueden resolverse en varias formas. Si se multiplica la primera por 10 y la segunda por 4 se obtiene

$$40x_1 + 160x_2 = 1\,280,$$
$$40x_1 + 80x_2 = 800.$$

Al restar la segunda ecuación de la primera se obtiene

$$80x_2 = 480$$
$$x_2 = 6.$$

El valor de x_1 se determina sustituyendo $x_2 = 6$ en cualquiera de las ecuaciones (con ambas se obtendrá el mismo valor de x_1). Al sustituir en la primera ecuación se obtiene

$$4x_1 + (16)(6) = 128$$
$$4x_1 = 128 - 96 = 32$$
$$x_1 = 8.$$

Compruebe el lector que al sustituir $x_2 = 6$ en la ecuación $10x_1 + 20x_2 = 200$ se obtiene el mismo resultado.

El último punto extremo se determina resolviendo

$$15x_1 + 10x_2 = 220,$$
$$10x_1 + 20x_2 = 200.$$

en forma simultánea. No se presentan los detalles de este cálculo. El lector debe poder demostrar, con los métodos anteriores, que la solución simultánea en este caso es

$$x_1 = 12,$$
$$x_2 = 4.$$

Ya hemos identificado los cinco puntos extremos. El siguiente paso es sustituir los valores correspondientes de x_1 y x_2 en la función objetivo, y ver cuál produce la ganancia máxima. La función objetivo es $115x_1 + 90x_2$.

Punto extremo	Valor de la función objetivo
(0, 0)	(115)(0) + (90)(0) = 0
(0, 8)	(115)(0) + (90)(8) = 720
(14.666 . . . , 0)	(115)(14.666 . . .) + (90)(0) = 1 686.67
(8, 6)	(115)(8) + (90)(6) = 1 460
(12, 4)	(115)(12) + (90)(4) = 1 740

El valor máximo de la función objetivo es 1 740, y se alcanza en el punto (12, 4). Esto concuerda con los resultados de la tabla S1-1.

Por consiguiente, en esta sección hemos demostrado que un método para determinar la solución óptima a un problema de programación lineal es señalar todos los puntos extremos, sustituir sus valores en la función objetivo y escoger el que produzca la máxima función objetivo

en problemas de maximización, o el valor mínimo de la solución objetivo en problemas de minimización. A continuación se indica cómo puede identificarse con rapidez el punto extremo óptimo en forma gráfica.

Identificación de la solución óptima en forma directa con método gráfico

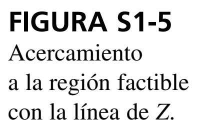
Se identifica directamente la solución óptima en la forma siguiente. La función objetivo es una combinación lineal de las variables de decisión. En nuestro ejemplo, la función objetivo es $115x_1 + 90x_2$. Veamos la familia de rectas definidas por la ecuación

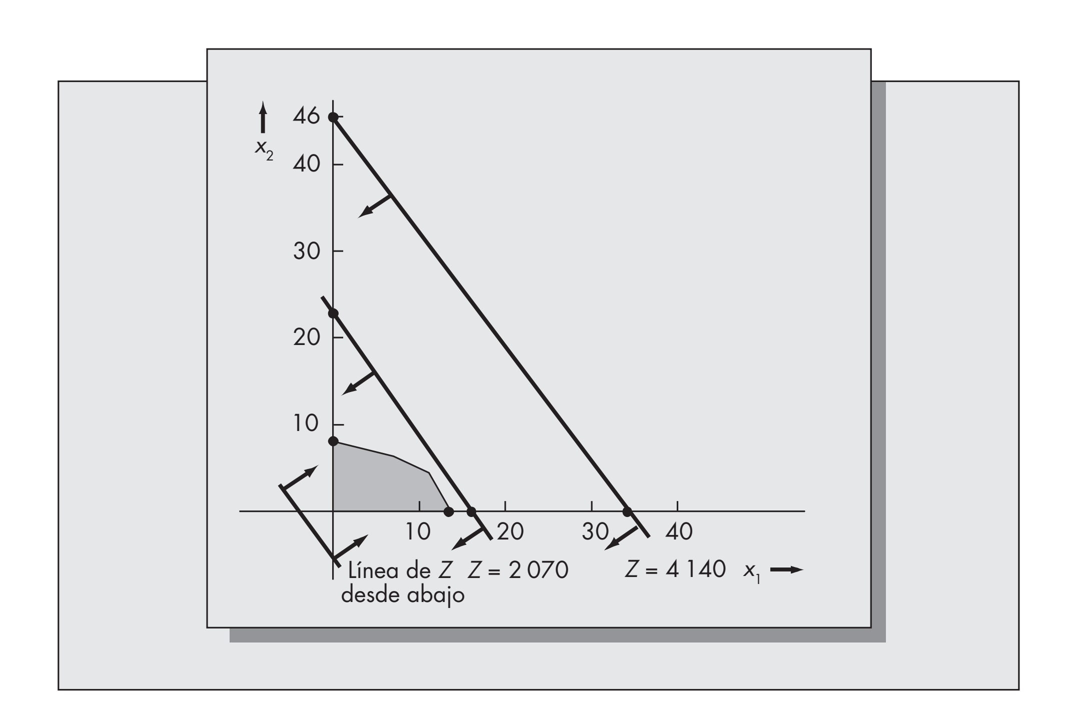
$$Z = 115x_1 + 90x_2.$$

Al variar Z se genera una familia de rectas paralelas. La variable Z es la ganancia que se obtiene al producir x_1 escritorios de cierre y x_2 escritorios normales, de tal manera que (x_1, x_2) está en la recta $Z = 115x_1 + 90x_2$. Por ejemplo, consideremos que $Z = 4\ 140$. La figura S1-5 muestra la recta $4\ 140 = 115x_1 + 90x_2$. Vemos que está totalmente fuera de la región factible, indicando que ninguna combinación factible de x_1 y x_2 da como resultado una ganancia de \$4 140. Al reducir Z a 2 070 baja la línea Z y se acerca a la región factible, lo cual también se ve en la figura S1-5.

El método gráfico para identificar la solución óptima es elegir un valor de Z, como por ejemplo $Z = 3\ 000$, que nos lleve fuera de la región factible, colocar una regla en la línea Z y mover la regla, manteniéndola paralela a Z, hacia la región factible. El punto extremo al que se llega primero es la solución óptima. Una vez determinado gráficamente cuál punto extremo es el óptimo, se pueden obtener sus coordenadas despejando las ecuaciones correspondientes, como ya se vio. Este método evita tener que identificar todos los puntos extremos de la región factible.

Sin embargo, puede surgir un problema. Si se elige un valor inicial pequeño de Z y se mueve hacia la región factible desde abajo, con este método se identificará un punto extremo *diferente*. Por ejemplo, supóngase que en la figura S1-5 se elige $Z = -3\ 000$. Entonces la línea Z estaría bajo la región factible. Al trasladar la línea Z hacia la región factible, el primer punto extremo que se encuentra es el origen. Esto significa que el origen resuelve el ejemplo S1.1 con


FIGURA S1-5
Acercamiento a la región factible con la línea de *Z*.



un objetivo de minimización, no de maximización (es decir, que si lo que se desea es *minimizar* la ganancia, la mejor estrategia es no producir escritorios).

Por lo tanto, con este método se identifican dos puntos extremos: uno que corresponde a la solución máxima y otro a la mínima. En este caso es obvio cuál es cuál. Cuando no es obvio, se debe determinar el valor (x_1, x_2) para cada punto extremo y sustituirlo en la función objetivo, para asegurar cuál es la solución mínima y cuál es la máxima.

S1.5 EL MÉTODO SIMPLEX: UNA PERSPECTIVA

El método simplex es un algoritmo que va en forma secuencial de un punto extremo a otro hasta llegar a la solución óptima. Si el origen (esto es, todas las variables del problema igualadas a cero) es una solución factible, sirve como el punto extremo inicial. En cada iteración, el método considera todos los puntos extremos adyacentes (los que se pueden alcanzar moviéndose a lo largo de un borde), y llega a aquel que produce la máxima mejora en la función objetivo. El algoritmo continúa moviéndose de un punto extremo a otro adyacente y finalmente termina cuando se llega a la solución óptima.

Para el problema del ejemplo S1.1, el origen es factible, de modo que es la solución factible inicial. Los dos puntos extremos adyacentes al origen son $(x_1, x_2) = (14.6667, 0)$ y $(x_1, x_2) = (0, 8)$. La máxima mejora en la función objetivo se obtiene pasando al punto (14.6667, 0). Hay dos puntos extremos adyacentes a (14.6667, 0), que son (0, 0) y (12, 4). Es claro que se obtiene una mejora mayor pasando a (12, 4). En este punto el método reconoce que la solución actual es la óptima, porque un movimiento a otro punto extremo adyacente disminuye la ganancia.

En el peor de los casos, el método simplex necesitaría buscar en todos los puntos extremos de la región factible antes de identificar la solución óptima. Si esto fuera lo normal, el método simplex no sería una solución práctica. Veamos por qué.

Suponga que un programa lineal tiene 25 variables y 25 restricciones menores o iguales. Para cada restricción se agrega una variable de holgura que convierte el problema en uno con restricciones de igualdad. A esto se le llama forma estándar. Por consiguiente, en la forma estándar se tiene un problema con 50 variables y 25 restricciones. Cada solución básica (punto extremo) corresponde a igualar 25 variables a cero y resolver el sistema resultante de 25 ecuaciones lineales con 25 incógnitas. En consecuencia, la cantidad total de soluciones (esto es, de puntos extremos) es igual a la cantidad de combinaciones de 50 objetos tomando 25 a la vez. Sucede que esto es aproximadamente 1.264×10^{14} (alrededor de 126 billones). Para que el lector tenga una idea de la magnitud de este número, suponga que se cuenta con un programa de cómputo que pudiera identificar 100 puntos extremos por segundo. Con esa rapidez se necesitarían ¡unos 40 000 años para calcular todos los puntos extremos en un problema de este tamaño!

En realidad es una fortuna que el método simplex casi nunca necesite buscar todos los puntos extremos de la región factible para descubrir la solución óptima. De hecho, para un problema de este tamaño y en promedio, se requiere evaluar tan sólo unos 25 puntos extremos.[4] Por consiguiente, el método simplex es un procedimiento muy eficiente de solución.

---

[3] Cuando el origen no es factible, hay técnicas para determinar una solución inicial factible con la cual iniciar el método.

[4] Sólo recientemente se comprendió la razón de ello. Su demostración requiere matemáticas muy sofisticadas.

No investigaremos la mecánica del método, ni la teoría en la que se basan sus conceptos. Con la facilidad actual para acceder al cómputo, y la gran disponibilidad de excelentes programas, es difícil que alguien que tenga un problema real lo resuelva en forma manual. Hay muchos y excelentes textos que penetran con más profundidad en los aspectos teóricos y computacionales de la programación lineal. Un buen punto de partida para el lector interesado es el de Hiller y Lieberman (1990). Una exposición más detallada de la teoría se puede encontrar en el libro de Hadley (1962).

S1.6 RESOLUCIÓN DE PROBLEMAS DE PROGRAMACIÓN LINEAL CON EXCEL

Excel se ha convertido en el líder del mercado en programas de hoja de cálculo. Las versiones actuales de Excel incluyen al Solver, que puede manejar problemas de programación lineal y no lineal. Aunque la confiabilidad de las porciones enteras y no lineales del programa resulte sospechosa, la parte de programación lineal de Solver es excelente.

Debido a que Solver es parte de un programa de hoja de cálculo, los problemas no se introducen algebraicamente (como en el sistema LINDO, analizado en las ediciones anteriores de este libro). Considere el ejemplo S1.1. La representación algebraica de ese problema es

$$\text{Maximizar } 115x_1 + 90x_2,$$

sujeta a

$$\begin{aligned} 10x_1 + 20x_2 &\le 200, \\ 4x_1 + 16x_2 &\le 128, \\ 15x_1 + 10x_2 &\le 220, \\ x_1, x_2 &\ge 0. \end{aligned}$$

Conviene escribir el problema en formato de matriz antes de teclear la información en la hoja de cálculo de Excel. La representación matricial para este problema es

Nombres de variables:	x_1	x_2	Operador	LD
Función objetivo:	115	90	max	
sujeta a				
Restricción 1	10	20	≤	200
Restricción 2	4	16	≤	128
Restricción 3	15	10	≤	220

La hoja de cálculo se verá muy parecida. Las únicas diferencias son que deben especificarse los lugares de los valores variables (que se recomienda estén en un renglón justo debajo de los nombres de las variables) y las fórmulas algebraicas para la función objetivo y las restricciones. Éstas se localizarán en una columna entre "Operador" y "LD". A esta columna la titularemos "Value" (valor) en la hoja de cálculo.

Obsérvese que no se necesita la columna titulada "Operador". Ayuda al usuario a mantener registro de la dirección de la función objetivo y las restricciones. Excel requiere esta información, pero se introduce manualmente cuando se presiona la opción Solver.

En este punto, la hoja de cálculo se debe parecer a la de la figura S1-6. Observe el renglón adicional titulado "Var Values" (valores de las variables) y la columna adicional titulada "Value" (valor). Es en esa columna donde se introduce la forma algebraica del problema de

FIGURA S1-6
Hoja de cálculo de Excel para el problema de Sidneyville.

Microsoft Excel - LINPROB3.XLS

File Edit View Insert Format Tools Data Window Help

profit =+C5*$C$4+D5*$D$4

	A	B	C	D	E	F	G
1			Sidneyville Problem				
2							
3	Var Names		x1	x2		Value	RHS
4	Var Values		0	0			
5	Objective Fn		115	90	Max	0	
6	st						
7	Constraint 1		10	20	<=	0	200
8	Constraint 2		4	16	<=	0	128
9	Constraint 3		15	10	<=	0	220
10							
11							
12							
13							
14							
15							
16							

Sensitivity Report 1 / Limits Report 1 / Sheet1 / Shee

programación lineal. Los lugares de los valores de la variable son las celdas C4 y D4. La forma algebraica de la función objetivo se capturará en la celda F5, y las restricciones en las celdas F7, F8 y F9. La fórmula para la celda F5 es =C5*$C$4+D5*$D$4. La fórmula se puede teclear o introducir con el ratón para señalar y hacer clic en las celdas.

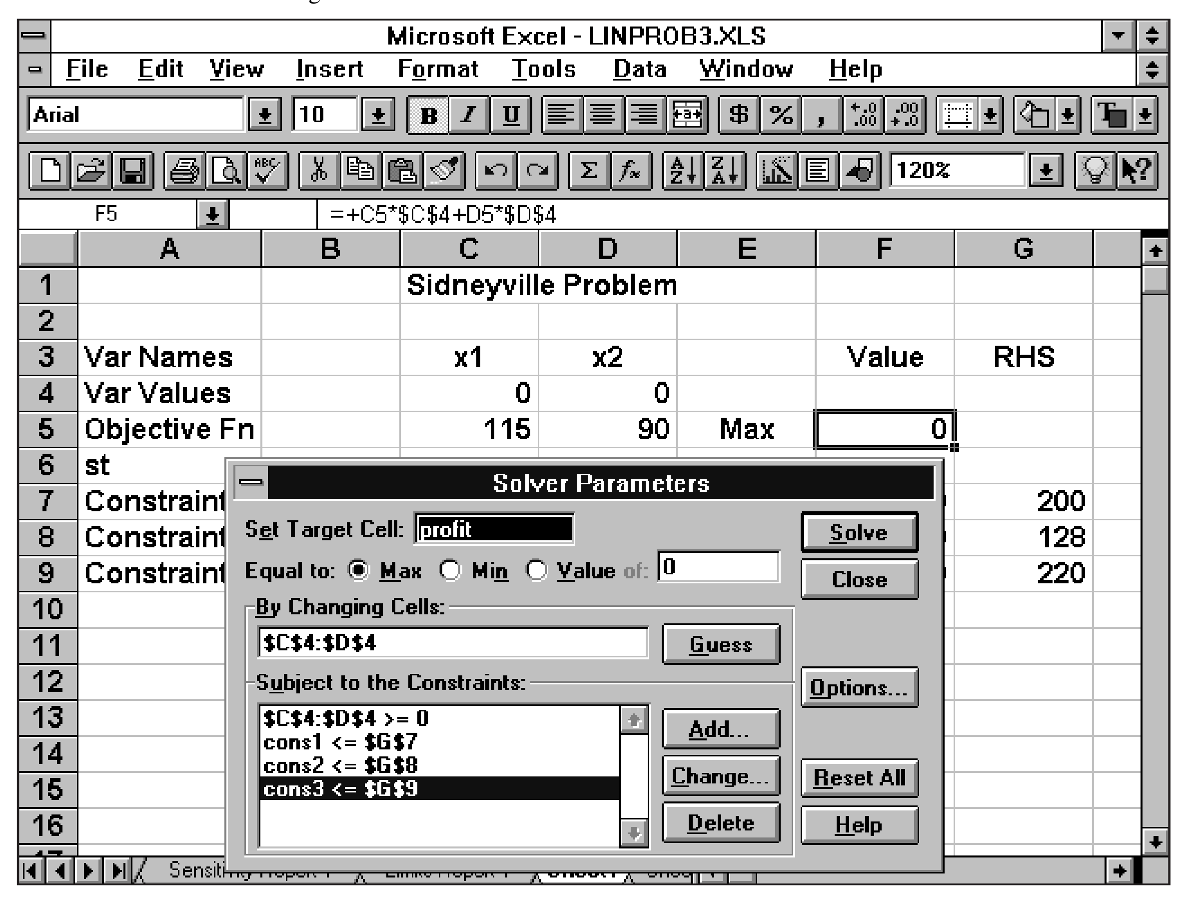
Observe que se ha usado el direccionamiento absoluto para los valores de variables (C4 y D4). Esto permite copiar la fórmula de la celda F5 en las celdas F7, F8 y F9 sin tener que volver a teclear la forma algebraica de las restricciones. También se pueden asignar etiquetas de nombre a esas celdas, para que después puedan referenciarse por nombre más que por localización de celda. (Esto se hace con más comodidad invocando la barra de fórmulas, entrando al área de etiquetas y tecleando el nombre que usted elija. Observe que la etiqueta “profit” (utilidad) aparece justo abajo del nombre del cuadro actual. Fue el primer nombre asignado a la celda F5.)

Compruebe que la forma algebraica de las restricciones sea correcta, después de copiar la celda F5 en las celdas F7, F8 y F9. Por ejemplo, la fórmula que corresponde a la celda F7 debe ser: =C7*$C$4+D7*$D$4.

Ponga ceros en las celdas correspondientes a los valores variables (C4 y C5). Después de que Solver completa la búsqueda, los valores óptimos aparecerán en esas celdas.

El problema está ya totalmente definido. Recurra a Solver para obtener la solución. Esto se hace a través de Herramientas, en la barra de menú, y presionando la opción Solver. El cuadro

FIGURA S1-7
Hoja de cálculo con el cuadro de diálogo Solver.



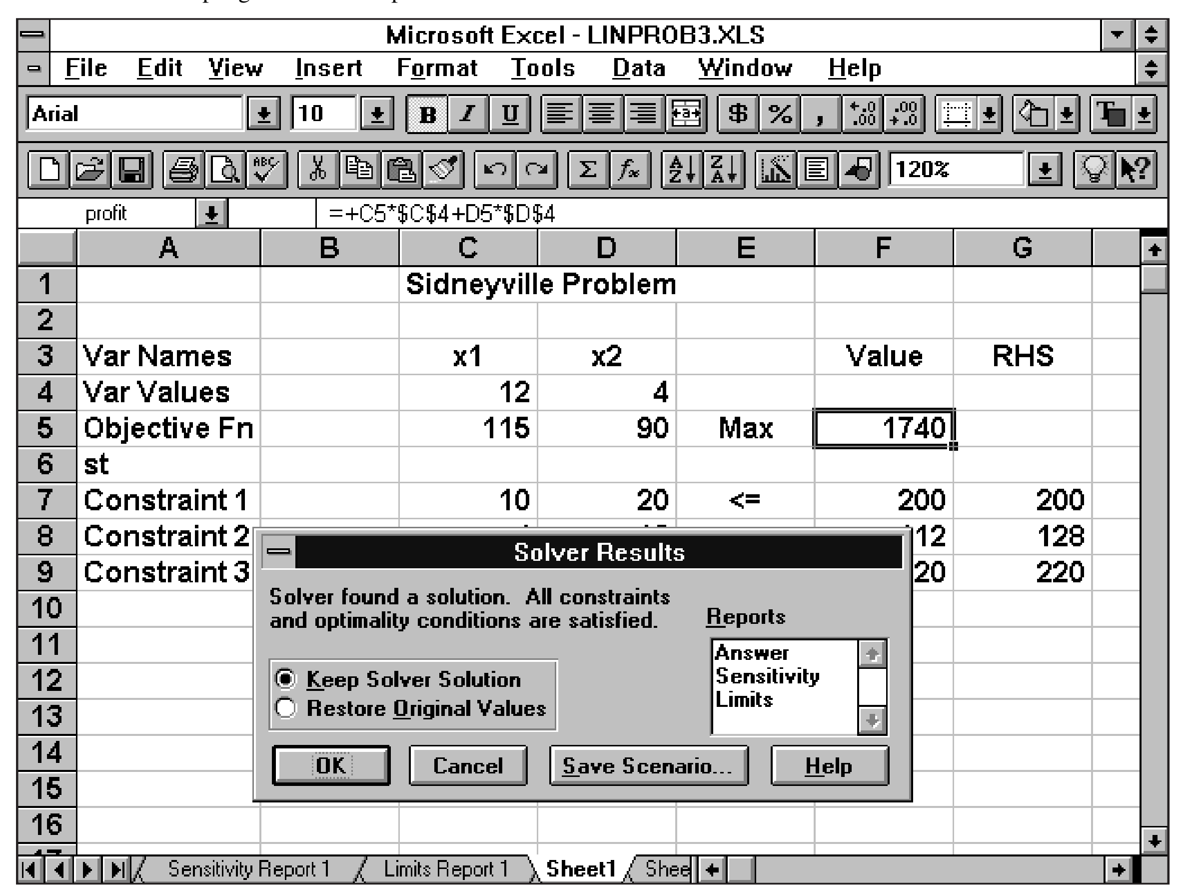
de diálogo de Solver aparecerá como en la figura S1-7. Primero hay que establecer la ubicación de la celda meta. Esto corresponde a la ubicación de la celda de la fórmula algebraica para la celda de función objetivo F5 en nuestra hoja de cálculo. Se necesita presionar la tecla F5 o haciendo clic en la celda con el ratón. Observe que en este caso se asignó el nombre "profit" (utilidad) a la celda F5, así que simplemente teclee "profit" en la celda meta.

A continuación se especifica si el problema es de min (minimizar) o de max (maximizar). Excel llama las variables en el problema "changing cells" (celdas cambiantes). Usted debe decir a Solver dónde encontrar esas celdas. En esta hoja de cálculo son C4 y D4. (Se pueden indicar señalándolas con el ratón y haciendo clic, tecleando los lugares de celda, o tecleando nombres preasignados de celdas como se hizo para la función objetivo.)

Enseguida se debe indicar a Excel dónde encontrar las definiciones algebraicas de las restricciones. Éstas se incorporan al sistema una por una, haciendo clic en el botón "Add" (agregar). Para cada restricción primero se le indica al sistema dónde encontrar la forma algebraica del lado izquierdo (F7, F8 y F9 en nuestro caso), el operador lógico para la restricción (≤, = o ≥) y la ubicación del valor de la derecha (G7, G8 y G9 en nuestro caso).

Debido a que Solver es una herramienta de programación matemática, es necesario incluir dos piezas adicionales de información. Ambas pueden hacerse al mismo tiempo haciendo clic en la tecla "Options" (opciones) y eligiendo "Assume linear model" y "Assumme non-negative" (para suponer un modelo lineal y agregar la restricción de no negatividad). De esta forma,

FIGURA S1-8
Hoja de Excel donde se despliega la solución óptima.

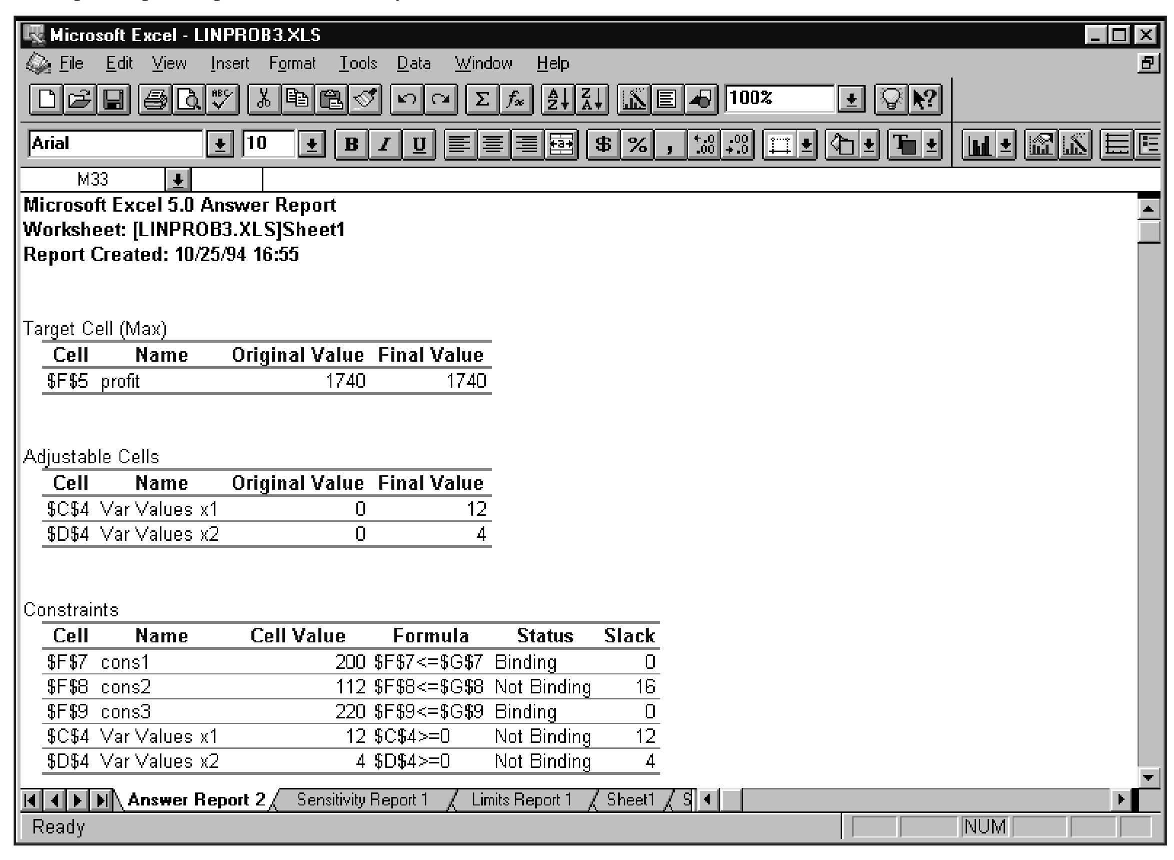


no es necesario que las restricciones de no negatividad se ingresen de forma explícita. (Nota: La opción de no negatividad no estaba disponible en la versión 5 de Excel o anteriores. Las restricciones de no negatividad tienen que ingresarse explícitamente en el problema.) Al decirle al programa que asuma un modelo lineal nos aseguramos de que se resuelva el problema por medio del algoritmo simplex más que el método de gradiente utilizado por los programas no lineales. Si su salida incluye valores de variables de Lagrange, sabrá que olvidó especificar esta opción.

En este punto, su cuadro de diálogo deberá verse como el de la figura S1-7. Observe que algunas de las celdas han recibido nombres, y que los nombres aparecen en el cuadro de diálogo y no en los lugares de las celdas. Si se usan celdas con nombres que tengan algún significado acerca del problema, se facilita después obtener la solución y los informes de sensibilidad. En la caja de diálogo se ha nombrado “profit” a la función objetivo (F5), y a las celdas que corresponden a las restricciones “constl”, “const2” y “const3”.

Compruebe haber capturado toda la información en forma correcta y haber especificado un problema lineal. Ahora tan sólo haga clic en el botón (Solve) y Excel entrará en acción y producirá una solución rápidamente. Después, la hoja de cálculo debe verse como la de la figura S1-8. Observe que los valores de las variables en las celdas C4 y D4 reflejan ahora la solución óptima 12 y 4, respectivamente. El valor en la celda F5 es $1 740, que es la ganancia óptima. Los de las celdas C7, C8 y C9 son los valores de los lados izquierdos de las restricciones.

FIGURA S1-9
Informe de respuesta para el problema de Sidneyville.



Microsoft Excel - LINPROB3.XLS

Microsoft Excel 5.0 Answer Report
Worksheet: [LINPROB3.XLS]Sheet1
Report Created: 10/25/94 16:55

Target Cell (Max)

Cell	Name	Original Value	Final Value
$F$5	profit	1740	1740

Adjustable Cells

Cell	Name	Original Value	Final Value
$C$4	Var Values x1	0	12
$D$4	Var Values x2	0	4

Constraints

Cell	Name	Cell Value	Formula	Status	Slack
$F$7	cons1	200	$F$7<=$G$7	Binding	0
$F$8	cons2	112	$F$8<=$G$8	Not Binding	16
$F$9	cons3	220	$F$9<=$G$9	Binding	0
$C$4	Var Values x1	12	$C$4>=0	Not Binding	12
$D$4	Var Values x2	4	$D$4>=0	Not Binding	4

Answer Report 2 / Sensitivity Report 1 / Limits Report 1 / Sheet1

Ready NUM

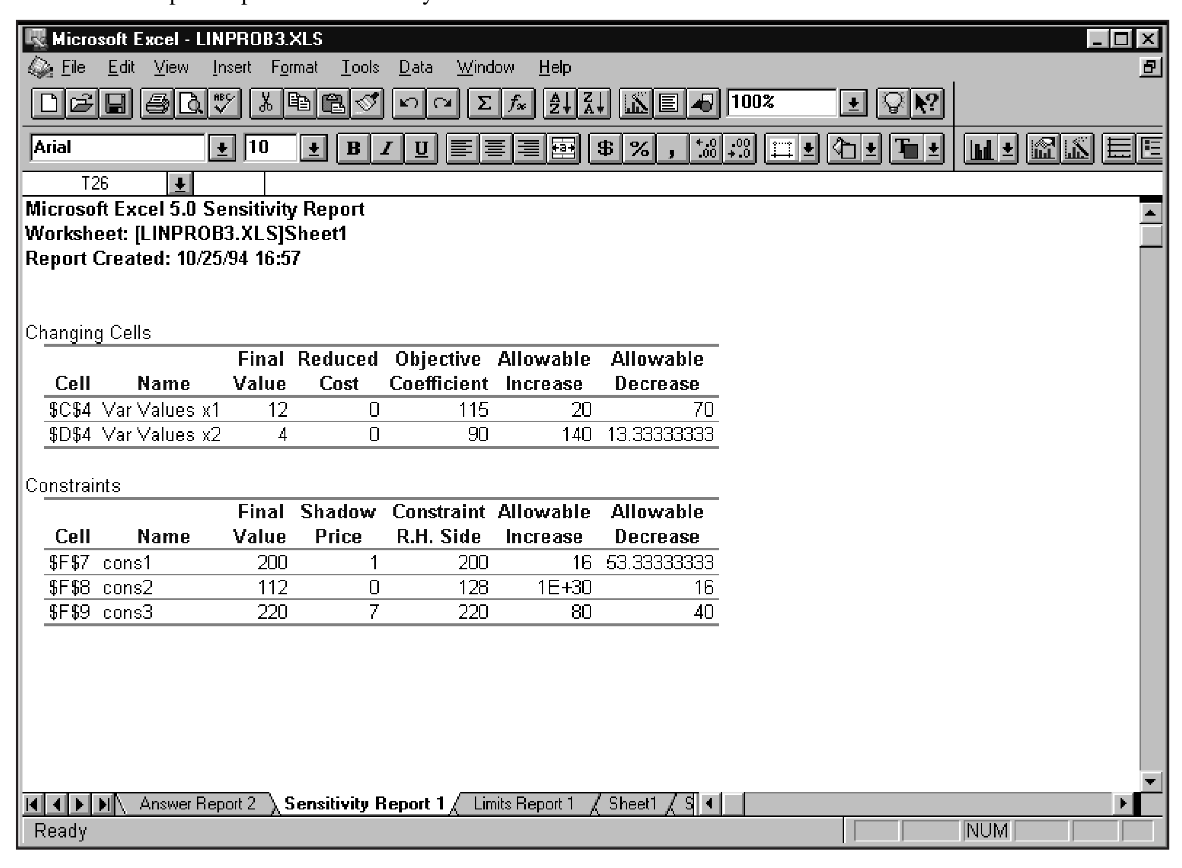
Aunque los valores óptimos de las variables y el de la función objetivo aparecen ahora en la hoja de cálculo, Excel presenta la opción de imprimir varias clases de informes. Los que importan en programación lineal son los informes de respuesta y de sensibilidad. Aparecen en distintas hojas del libro de trabajo y se ven en las figuras S1-9 y S1-10. La mayor parte de la información en el informe de respuesta aparece en la hoja de cálculo original. También se indicó cuáles restricciones son activas. Una variable que se menciona como "not binding" (no activa) es una que no es cero.

Captura eficiente de problemas grandes

Los pasos descritos para resolver problemas de programación lineal en una hoja de cálculo son para problemas pequeños. Sin embargo, Excel tiene varias propiedades que permiten una captura más eficiente de problemas mayores.

Una de ellas es la función SUMPRODUCT, que consiste en un vector o producto de arreglos que multiplica los elementos de un arreglo por los de otro, término a término, y suma los resultados. Esto significa que las formas algebraicas de la función objetivo y de las restricciones se pueden capturar con esta función. Recuérdese que para la celda F5 se usó la fórmula =C5*$C$4+D5*$D$4. También se pudo haber capturado como =SUMPRODUCT($C$4:$D$4, C5:D5). Si bien esto quizás no parezca una gran mejora, evita un tecleo excesivo cuando se capturan problemas grandes.

FIGURA S1-10
Informe de sensibilidad para el problema de Sidneyville.



Microsoft Excel - LINPROB3.XLS

File Edit View Insert Format Tools Data Window Help

Microsoft Excel 5.0 Sensitivity Report
Worksheet: [LINPROB3.XLS]Sheet1
Report Created: 10/25/94 16:57

Changing Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$C$4	Var Values x1	12	0	115	20	70
$D$4	Var Values x2	4	0	90	140	13.33333333

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$F$7	cons1	200	1	200	16	53.33333333
$F$8	cons2	112	0	128	1E+30	16
$F$9	cons3	220	7	220	80	40

Answer Report 2 | Sensitivity Report 1 | Limits Report 1 | Sheet1

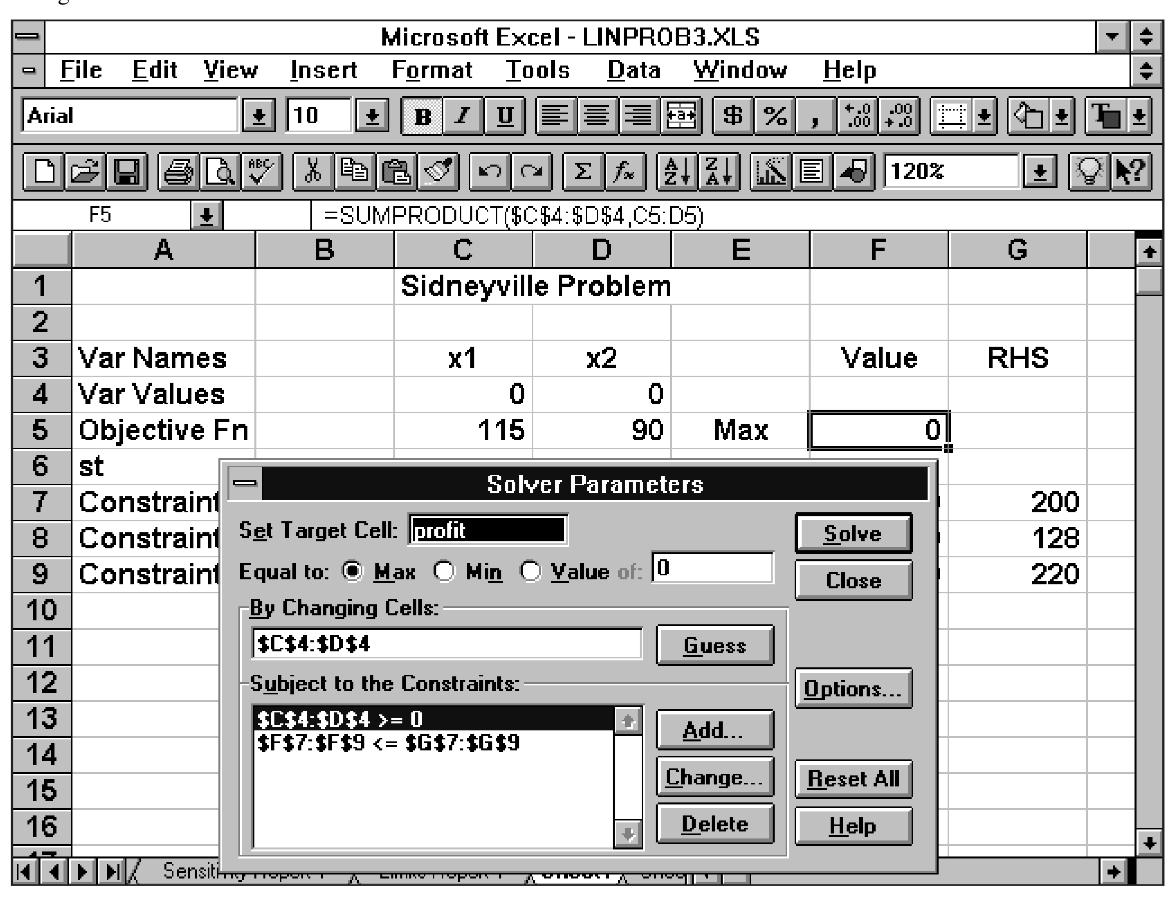
Ready NUM

El segundo atajo para capturar problemas grandes es agrupar las restricciones por tipo, y capturar todas en un grupo a la vez. En el problema de Sidneyville hay tres restricciones ≤. Éstas pudieron capturarse con un comando semejante al usado para introducir las restricciones de no negatividad. Las fórmulas adecuadas de las restricciones aparecen en las celdas F7, F8 y F9. El comando único para capturarlas en Solver es: $F$7:$F$9 ≤ $G$7:$G$9. Esto se puede teclear directamente, o capturarlo señalando las celdas correspondientes en la hoja de cálculo. En la figura S1-11 se ve el cuadro de diálogo para el problema de Sidneyville al aplicar este método.

Al usar SUMPRODUCT y teclear las restricciones en grupos se puede ahorrar mucho tiempo en los problemas grandes. Sin embargo, el cuadro de diálogo de Solver no es tan informativo, porque sólo aparecen las ubicaciones de las celdas y no sus nombres. Para problemas con menos de 10 restricciones en total, se aconseja teclear las restricciones una por una.

Antes se hizo notar que una ventaja de la hoja de cálculo en la solución de programas lineales es que puede formar una plantilla de propósito general. Una plantilla podría tener hasta 10 variables y 10 restricciones. Los nombres de variables podrían ser x1 hasta x10, y los de las restricciones constl hasta constl0. Las funciones SUMPRODUCT para la función objetivo y las restricciones se programarían de antemano. A continuación sólo se teclearían los coeficientes del problema por resolver y se guardarían con un nuevo nombre de archivo. (Es mucho más rápido que teclear la representación algebraica completa para cada problema, que es lo que se hace cuando se usa otro sistema de programación lineal.)

FIGURA S1-11
Caja de diálogo de Solver con introducción eficiente de los datos.



S1.7 INTERPRETACIÓN DEL INFORME DE SENSIBILIDAD

Precios sombra

Una cuestión interesante es el valor de recursos adicionales en un problema de programación lineal. En el ejemplo S1.1, los recursos son los tres tipos de madera: pino, cedro y arce. Un incremento en el nivel de alguno de estos recursos da como resultado un incremento en el valor del lado derecho de la restricción apropiada. El informe de sensibilidad (figura S1-10) proporciona información acerca del valor a la función objetivo de recursos adicionales. Tal información está contenida en los valores de los precios sombra para las restricciones.

El precio sombra se define como una mejora en la función objetivo realizada agregando una unidad adicional de un recurso. Para el ejemplo S1.1, la primera restricción se refiere a la cantidad de pino necesaria. Debido a que esta restricción es obligatoria en la solución óptima (porque el valor final y los valores de restricción de lado derecho son los mismos), es probable que si contamos con pino adicional podamos incrementar los ingresos. El precio sombra nos dice en que es benéfico. Debido a que el precio sombra para esta restricción es $1, significa que por cada metro cuadrado unitario adicional de pino, el objetivo "profit" (utilidad) se incremente en $1. (Esto sólo se mantendrá dentro de un cierto rango, como se describe a continuación.) Considere la segunda restricción (cedro). Debido a que los valores finales del

lado derecho para esta restricción son diferentes (112 contra 120), existe un descuido en esta restricción, lo cual significa que no estamos consumiendo toda la cantidad de cedro en la solución óptima, y el cedro adicional no mejorará la utilidad. Esto obedece al que el precio sombra de esta restricción es cero. El precio sombra final de 7 indica que cualquier metro cuadrado adicional de arce contribuye con $7 a la utilidad.

Lados derechos y coeficientes de la función objetivo

Los precios sombra seguirán siendo válidos siempre que no cambie la base óptima. (La base es el conjunto de variables positivas en la solución final.) Las columnas "Allowable Increase" y "Allowable Decrease" (incremento permisible y decremento permisible) indican que en ese rango los precios sombra permanecerán como válidos. Esto significa que podemos determinar el efecto en la función objetivo de los cambios en los lados derechos de la restricción sin resolver el problema.

La primera parte del informe de sensibilidad (figura S1-10) da los valores de los coeficientes de función objetiva para los cuales los precios sombra permanecen como válidos. Los valores actuales de las utilidades son 115 y 90 dólares, respectivamente. Los precios muestra son válidos siempre que el primer coeficiente de función objetiva no se incremente más allá de 20 o disminuya más de 70 (es decir, el coeficiente de la función objetivo para x_1 está entre 45 y 135). De forma similar, el rango permisible para el coeficiente de la función objetivo de x_2 es de $76\frac{2}{3}$ a 230.

La segunda parte del informe de sensibilidad da los valores de los lados derechos de las restricciones para los que permanecen como válidos los precios sombra. Por lo tanto, el precio sombra de 1 dólar es válido para un incremento de 16 o menos y un decremento de $53\frac{1}{3}$ o menos en el lado derecho de la restricción. Esto es, el lado derecho de la primera restricción puede ser cualquier valor entre $146\frac{2}{3}$ y 216. El precio sombra de 0 para el cedro es válido en cualquier incremento (IE+30 debe interpretarse como infinito) y un decremento de 16 o menos, y el precio sombra para el arce es válido para un incremento de 80 y un decremento de 40 o menos.

Si los coeficientes de la función objetivo o los lados derechos se incrementan o disminuyen más allá de los rangos permisibles, los precios sombra dejarán de ser válidos y el problema tendrá que resolverse para determinar el efecto. Observe que Excel usa la norma de que un precio sombra positivo significa un incremento en la función objetivo por incremento de unidad en el lado derecho de la restricción; un precio sombra negativo significa una disminución de la función objetivo por incremento de unidad en el lado derecho, sin importar si el problema es max o min. (Otros sistemas de programación lineal pueden tener otras normas.) Observemos también que los cambios a los coeficientes de función objetivo o lados derechos sólo se pueden realizar uno a la vez en estos rangos. Aquí no comentaremos las reglas a cambios simultáneos en los lados derechos o los coeficientes de la función objetivo, que son mucho más complejos.

Agregar una nueva variable

Podemos usar los resultados del análisis de sensibilidad para determinar si es rentable agregar una nueva actividad (variable) sin resolver el problema. En el caso de problemas pequeños, como en el ejemplo S1.1, la introducción y solución del nuevo problema en la computadora es un asunto sencillo y rápido. Sin embargo, en las aplicaciones reales, el número de variables de decisión y restricciones puede resultar en cientos o incluso miles. Reingresar un problema de esta magnitud es una tarea titánica que debe evitarse siempre que sea posible.

Supongamos que la compañía está produciendo un tercer producto, un tocador, que requerirá de las mismas maderas que se utilizan para hacer los escritorios. Cada tocador proporcionará una utilidad de $75, pero cada uno requerirá de 8 metros cuadrados de pino, 6 de cedro y 10 de arce. Podemos determinar si vale la pena producir los tocadores además de los escritorios resolviendo el problema con las tres actividades y comparando los valores de las funciones objetivo en las soluciones óptimas.

Existe una forma más rápida. Los precios duales que se presentan en la figura S1-10 nos dicen el valor de cada unidad de recurso en la solución actual. El decremento en utilidad que resulta de reducir el abastecimiento de pino es de $1 por metro cuadrado, que se traduce en $8 por los 8 metros cuadrados de pino. No hay ningún costo por disminuir el abastecimiento de cedro. El costo por disminuir el abastecimiento de arce en 10 metros cuadrados es (10)(7) = $70. Por lo tanto, la disminución total en utilidades proveniente del consumo de recursos que implica producir un tocador es de $78. La contribución a la utilidad es de sólo $75. Concluimos que no es óptimo producir tocadores además de escritorios con los recursos que se tienen actualmente. Determinamos, sin embargo, que si hubiese sido rentable producir el tocador, tendríamos que resolver el problema de las tres actividades para encontrar el número óptimo de escritorios y tocadores que habría que producir.

Utilización del análisis de sensibilidad

Para un mejor entendimiento de la información que aparece en la figura S1-1, consideremos las siguientes preguntas:

Ejemplo *S1.1* *(continuación)*

a) El gerente de ventas de Sidneyville ha renegociado el contrato de escritorios normales y espera obtener una ganancia de 125 dólares por cada uno. Entusiasmado, comparte esta información con el gerente de producción de la compañía, esperando que la mezcla óptima de escritorios de cierre y normales genere un cambio. ¿Sucederá esto?

b) Supongamos que el nuevo contacto también tiene una mayor utilidad para los escritorios de cierre. Si la nueva ganancia para los escritorios de cierre es de $140, ¿cómo puede cambiar esto la solución óptima?

c) Una compañía maderera ha ofrecido vender a Sidneyville 50 metros cuadrados adicionales de arce a $5 cada uno. Con base en la función objetivo original, ¿recomendaría usted que se aceptara la oferta?

d) Suponiendo que Sidneyville compra 50 metros cuadrados de arce, ¿cómo se ve afectada la solución óptima?

e) La compañía está considerando un escritorio de pino que requerirá 25 metros cuadrados de pino y ninguna otra madera. ¿Qué utilidad para los escritorios de pino se requerirá para que esta producción sea rentable, suponiendo los niveles actuales de recursos y utilidades en los escritorios normales y de cierre?

f) Durante la inspección, el departamento de calidad descubrió que 50 metros cuadrados de pino estaban dañados por el agua y no podían utilizarse. ¿Resultaría óptimo producir ambos escritorios bajo estas circunstancias? ¿Cambiaría la mezcla de producto?

Solución

a) De acuerdo con la figura S1-10, el incremento permisible en el coeficiente de la función objetivo para la variable x_2 (los escritorios normales) es 140. Debido a que el incremento a 125 dólares sigue estando dentro del rango permisible, la mezcla óptima de escritorios normales y de cierre permanecerá igual: a saber, $x_1 = 12$ y $x_2 = 4$.

b) El incremento permisible en la función objetivo para los escritorios de cierre es 20, o hasta un valor máximo de 135. Como 140 está fuera del rango permisible, es posible que la base cambie. Sin embargo, los rangos permisibles en la figura S1-10 son válidos únicamente si la utilidad para los escritorios normales es de $90. Los rangos permisibles cambiarán cuando la utilidad de los escritorios normales se modifique a 125 dólares, aunque no cambie la solución óptima. La salida del inciso *a*) (esto es, con utilidades de 115 y 125 dólares) es

OBJ COEFFICIENT RANGES

VARIABLE	CURRENT COEF	ALLOWABLE INCREASE	ALLOWABLE DECREASE
X1	115.000000	72.500000	52.500000
X2	125.000000	105.000000	48.333330

Esto muestra que el incremento permisible en el coeficiente ahora es de 72.5. Debido a que 140 se encuentra dentro del rango permisible, la solución para los incisos *a*) y *b*) será la misma; también es igual a nuestra solución original de $x_1 = 12$ y $x_2 = 14$.

c) Debido a que el precio dual para la tercera restricción correspondiente al arce es 7, es rentable comprar el arce a $5 el metro cuadrado. El incremento permisible de lado derecho sobre el cual se aplica este precio dual es 80, así que es rentable comprar los 50 metros cuadrados adicionales.

d) Debido a que el incremento de 50 se encuentra dentro del rango permisible de lado derecho, sabemos que no cambiará la base. Esto es, continuará siendo óptimo producir tanto escritorios normales como de cierre. Sin embargo, si el lado derecho cambia, los valores de las variables básicas también cambiarán. Debemos resolver el problema con el nuevo valor de lado derecho para determinar la solución actualizada. La solución es

```
                    LP OPTIMUM FOUND AT STEP 1

OBJECTIVE FUNCTION VALUE

             1)            2090.000000

       VARIABLE         VALUE              REDUCED COST
       X1               17.000000              .000000
       X2                1.500000              .000000
```

Para conservar la factibilidad, redondeamos x_2 a 1. (Sin embargo, esta *no* es la solución completa óptima. La solución completa óptima es $x_1 = 18$ y $x_2 = 0$ con una utilidad de $2 070, que se obtiene de Excel al identificar como variables enteras tanto a x_1 como a x_2. La solución subóptima de $x_1 = 17$ y $x_2 = 1$ genera una utilidad de $2 045.)

e) El precio dual del pino es $1 por metro cuadrado. Debido a que cada escritorio consume 25 metros cuadrados de pino, la utilidad de cada escritorio de pino debe exceder los $25 para que sea rentable su producción.

f) El lado derecho de la primera restricción puede disminuir tanto como 53.333330 y la base actual permanecerá como óptima. Eso significa que un decremento de 50 metros cuadrados no cambiará la base; continuará siendo rentable producir ambos escritorios. Sin embargo, disminuirán las cantidades de producción. Debemos resolver el problema para determinar los niveles correctos de las nuevas cantidades. Éstos son

```
LP OPTIMUM FOUND AT STEP 2
 OBJECTIVE FUNCTION VALUE
        1)     1690.000000
  VARIABLE     VALUE                REDUCED COST
        X1     14.500000                 .000000
        X2       .250000                 .000000
```

De nuevo, necesitamos redondear estas variables. El redondeo de x_1 y x_2 hacia abajo garantiza la factibilidad. Si producimos 14 escritorios de cierre y 0 escritorios normales, necesitaremos 140 metros cuadrados de pino (hay 150 disponibles), 56 metros cuadrados de cedro (hay 128 disponibles), y 210 metros cuadrados de arce (hay 220 disponibles). Parece que no se cuenta con la madera suficiente para producir un escritorio adicional de cualquier tipo, así que dejamos la solución en $x_1 = 14$ y $x_2 = 0$. (Esta es la solución entera óptima.)

S1.8 RECONOCIMIENTO DE PROBLEMAS ESPECIALES

Se pueden encontrar algunas dificultades al resolver problemas de programación lineal. En esta sección se describen sus causas y cómo se reconocen al usar Excel.

Soluciones no acotadas

La región factible de un programa lineal no necesariamente se acota. Para el problema de Sidneyville de la figura S1-4 sí es acotada. Sin embargo, considérese el siguiente problema de programación lineal.

Ejemplo *S1.2* Maximizar

$$2x_1 + 3x_2$$

sujeta a

$$x_1 + 4x_2 \geq 8$$

$$x_1 + x_2 \geq 5$$

$$2x_1 + x_2 \geq 7$$

$$x_1, x_2 \geq 0.$$

La figura S1-12 muestra la región factible. Observe que no está acotada. Como x_1 y x_2 pueden agrandarse tanto como se desee, no hay límite para el tamaño de la función objetivo. Cuando sucede esto, el problema es no acotado y no hay solución óptima.

En el caso de Excel, Solver escribe valores muy grandes de las variables del problema y presenta un mensaje indicando que los valores buscados no convergen. En la figura S1-13 aparece el resultado de Excel para este problema.

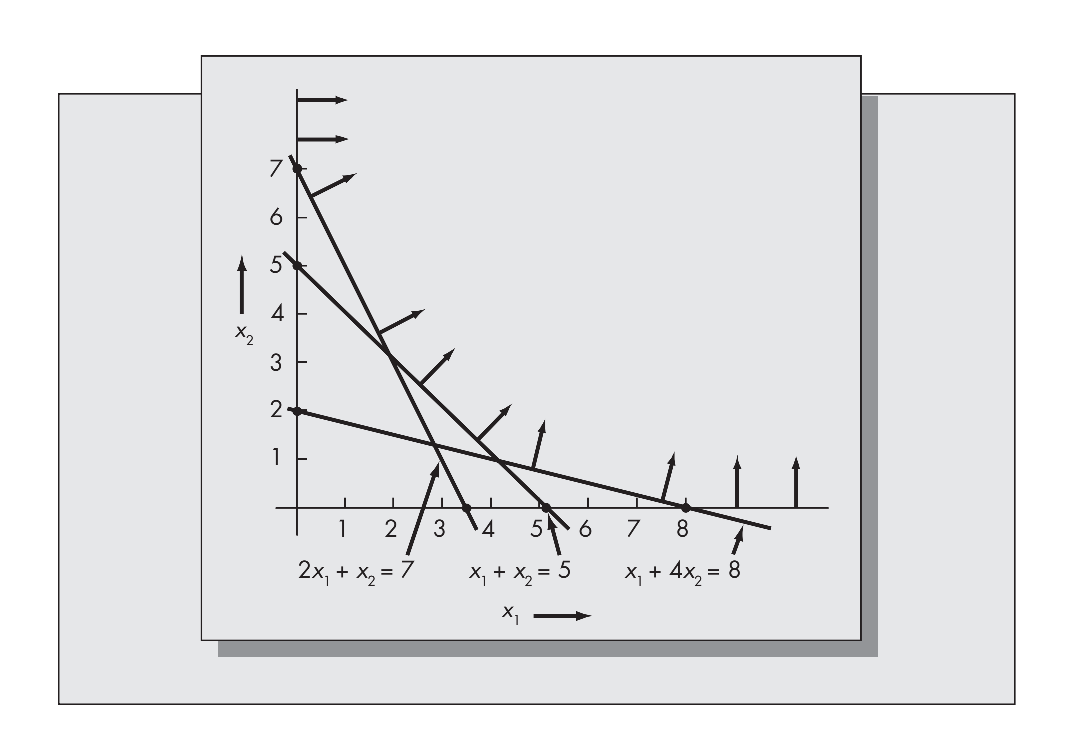
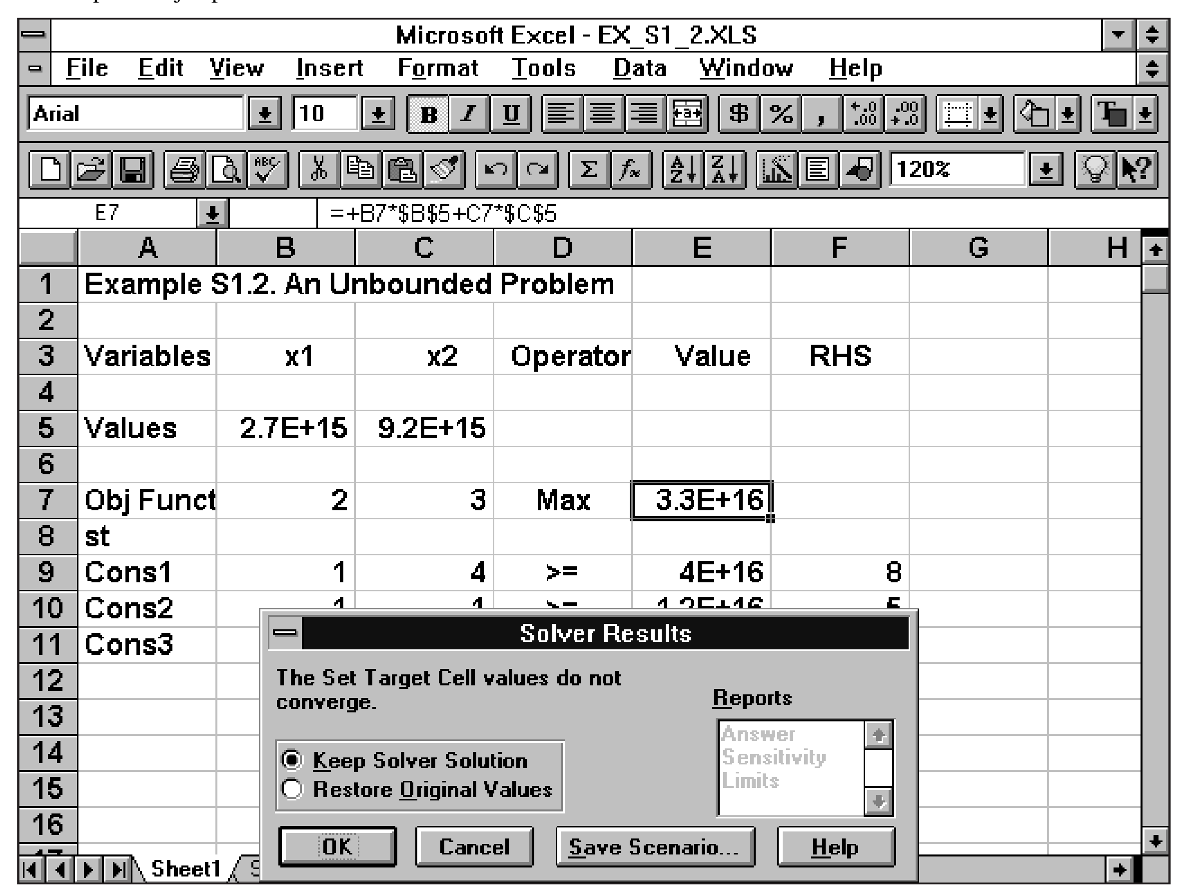


FIGURA S1-12
Región factible para el ejemplo S1.2.

FIGURA S1-13
Salida de Excel para el ejemplo S1.2.



Región factible vacía

Es posible que dos o más restricciones sean inconsistentes. Cuando eso suceda no habrá solución factible. Veamos el siguiente problema.

Ejemplo *S1.3* Maximizar

$$2x_1 + 3x_2$$

sujeta a

$$x_1 + 4x_2 \leq 8$$

$$x_1 + x_2 \geq 10$$

$$x_1, x_2 \geq 0.$$

La región factible para este ejemplo aparece en la figura S1-14. Observe que no hay intersección de los semiespacios definidos por las dos restricciones en el cuadrante positivo. En este caso, la región factible está vacía y se dice que el problema no es factible. De nuevo, el que aparezca una solución no significa que sea factible. La solución $x_1 = 8$ y $x_2 = 0$ no es factible, porque da como resultado una holgura negativa en el segundo renglón. El resultado con Excel se ilustra en la figura S1-15. Observe que la solución $x_1 = 8$ y $x_2 = 0$ que se muestra no es factible porque da como resultado una holgura negativa en la primera restricción.

FIGURA S1-14
Región factible para el ejemplo S1.3.

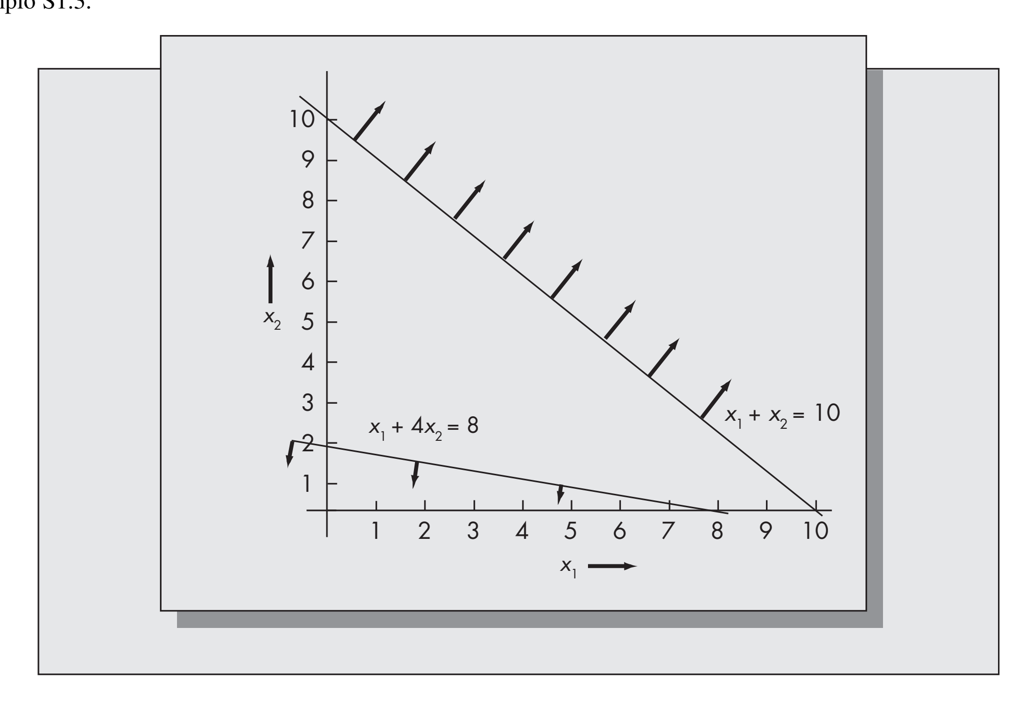
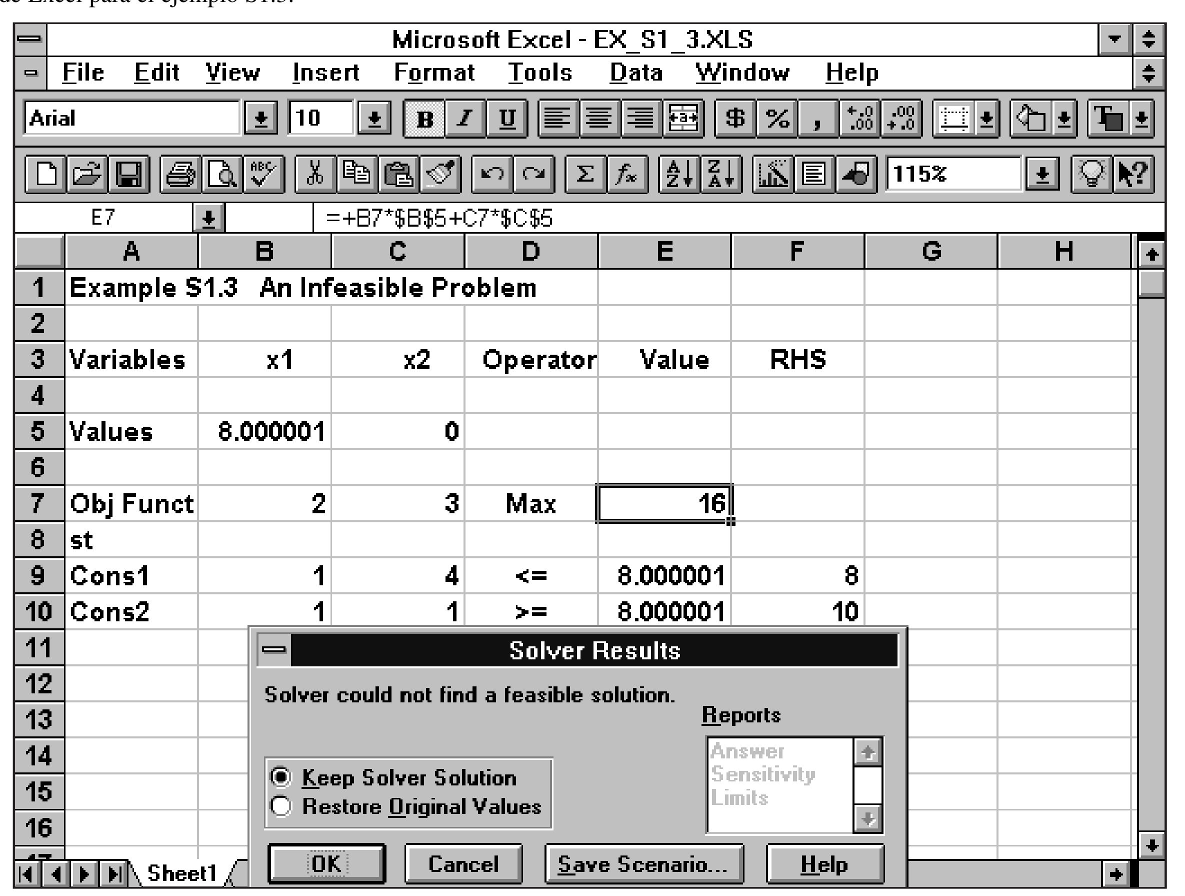


FIGURA S1-15
Salida de Excel para el ejemplo S1.3.



Degeneración

En programación lineal existen dos tipos de variables: las básicas y las no básicas. La cantidad de variables básicas es igual a la cantidad de restricciones, y esas variables pueden ser originales, de holgura o de exceso. ¿Qué define a las variables básicas? Véase cualquier programa lineal en su forma estándar. Al incluir las variables de holgura y de exceso, todas las restricciones quedan expresadas como ecuaciones. En la forma estándar siempre habrá más variables que restricciones. Suponga que después de haber expresado un problema de programación lineal en su forma estándar, hay $n + m$ variables y n restricciones. Una solución básica se determina igualando a cero m variables y resolviendo las n ecuaciones resultantes con n incógnitas. En la mayoría de los casos, los valores de las n variables básicas serán positivos. Se presenta una solución degenerada cuando una o más variables básicas son cero en la solución óptima. En Excel, una variable básica es aquella con costo reducido cero o precio dual cero. La degeneración se presenta cuando el valor de una variable es cero, y su costo reducido o precio dual también es cero.

¿Por qué nos interesa la degeneración? Es posible que si se presentan soluciones degeneradas, el método simplex entrará en ciclo a través de un conjunto de soluciones y nunca reconozca la solución óptima. El fenómeno de entrar en ciclos nunca se ha observado en la práctica, y la mayoría de los programas de cómputo tienen medios para garantizar que nunca sucederá. La degeneración es un asunto que no debe considerarse motivo de preocupación, pero al cual no debe perdérsele de vista.

Soluciones óptimas múltiples

La solución de un programa lineal no siempre es única. Hay casos en que existen varias soluciones. Ya se vio en este capítulo que se pueden resolver dos problemas variables con medios gráficos acercándose a la región factible con la línea Z. Suponiendo que se llegue a la región factible desde el lado correcto, el primer punto factible que toca la línea Z es la solución óptima.

Sin embargo, suponga que la línea Z es paralela a una de las restricciones. En ese caso no toca primero un solo punto, sino todo un borde.

Ejemplo *S1.4*

Véase la región factible que muestra la figura S1-12, que corresponde a las restricciones del ejemplo S1.2. Supóngase que la función objetivo es min $3x_1 = 3x_2$. Entonces la línea Z tiene una pendiente igual a -1 y es paralela a la frontera de restricción $x_1 + x_2 = 5$. Cuando la línea Z se acerca a la región factible, encuentra el borde definido por esta restricción y ambos puntos extremos en dicho borde. Esto significa que los dos puntos extremos y también todos los puntos a lo largo del borde son óptimos.

Desafortunadamente, Excel notifica que hay soluciones óptimas múltiples para este problema. La única pista es que la solución es degenerada; la variable de exceso para la tercera restricción tiene a la vez precio dual cero y valor cero. Nuestra solución gráfica indica que ambos puntos extremos, (2, 3) y (4, 1), son óptimos, y por consiguiente también lo son todos los puntos a lo largo del borde que une a esos puntos extremos. [Los puntos del borde se pueden representar con las siguientes ecuaciones:

$$x_1 = \alpha(2) + (1 - \alpha)(4),$$

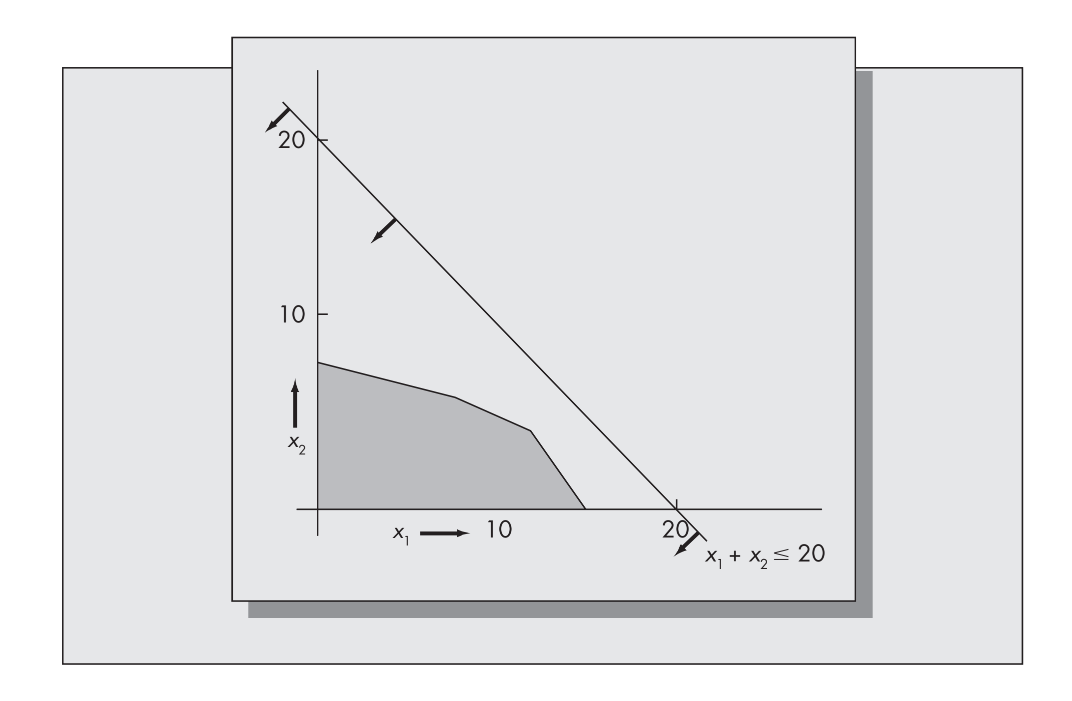
$$x_2 = \alpha(3) + (1 - \alpha)(1),$$

donde α es un número entre cero y uno. A esto se le llama combinación convexa de esos dos puntos extremos.]

Restricciones redundantes

Es posible que una o más restricciones sean redundantes. Esto significa que pueden eliminarse de la formulación sin afectar la solución. En problemas sencillos con dos variables se pueden reconocer gráficamente las restricciones redundantes, pues quedan fuera de la región factible. Excel no reconoce ni indica si una o más restricciones son redundantes. A veces esas restricciones pueden causar degeneración, pero también la degeneración puede originarse cuando las restricciones no son redundantes.

FIGURA S1-16
Región factible para el ejemplo S1.5 que muestra una restricción redundante.



Ejemplo *S1.5* Considere el ejemplo S1.1. Suponga que agregamos la siguiente restricción adicional:

$$x_1 + x_2 \leq 20.$$

La figura S1-16 muestra la región factible resultante. Es exactamente la misma que se ilustra en la figura S1-4. La restricción adicional no tuvo efecto, por estar totalmente fuera de la solución factible definida por las tres primeras restricciones. Naturalmente, la solución óptima será exactamente la misma.

Si se tuviera el problema formulado originalmente con las cuatro restricciones

$$10x_1 + 20x_2 \leq 200,$$

$$4x_1 + 16x_2 \leq 128,$$

$$15x_1 + 10x_2 \leq 220,$$

$$x_1 + x_2 \leq 20,$$

y se hubiera resuelto, el resultado indicaría que la solución óptima es $x_1 = 12$ y $x_2 = 4$, como antes, y no mostraría pista alguna para indicar que la última restricción es redundante. La única forma de verificarlo sería graficar la región factible, como se hizo en la figura S1-16. Sin embargo, esas gráficas sólo son posibles en problemas con dos variables.

¿Causa algún problema la redundancia? En realidad no. Claro que nos gustaría escribir nuestro programa lineal de la forma más económica posible, pero si una o más restricciones son redundantes, la solución óptima no se verá afectada.

S1.9 APLICACIÓN DE LA PROGRAMACIÓN LINEAL AL ANÁLISIS DE LA PRODUCCIÓN Y DE LAS OPERACIONES

En el capítulo 3 se indicó cómo se puede aplicar la programación lineal para calcular las soluciones óptimas (sujetas a errores de redondeo) en problemas de planeación agregada. Aunque es el único uso explícito de la programación lineal en este libro, han existido diversas aplica-

ciones en múltiples problemas de administración de operaciones.[5] Quizá la programación de producción y la distribución sean las dos áreas en las que las aplicaciones son más comunes.

Fisher y otros (1982) describen una aplicación de programación lineal al problema de proporcionar un programa coordinado de vehículos y un sistema de rutas para entregar productos consumibles a los clientes de la compañía Du Pont. El asunto principal fue determinar las rutas de entrega (ciclos) para los camiones que dan servicio a los clientes de la empresa en diversas regiones de Estados Unidos. Un camión refrigerador recorre un ciclo semanal que abarca varias docenas de clientes. La región más grande que se consideró (Chicago) tenía 16 ciclos y varios cientos de ciudades, mientras la región de Houston, la más pequeña, tenía 4 ciclos y menos de 80 ciudades.

La formulación matemática básica del problema fue un problema general de asignación. (El problema de asignación se describe en el capítulo 9. Es un problema de programación lineal en el cual las variables de decisión se restringen a ser ceros y unos.) La formulación matemática usada en este estudio es la siguiente:

1. Datos proporcionados

$$d_{ik} = \text{Costo de incluir al cliente } i \text{ en el ciclo } k,$$
$$a_i = \text{Demanda del cliente } i.$$

2. Variables del problema

$$y_{ik} = \begin{cases} 1 & \text{si el cliente } i \text{ es asignado al ciclo } k, \\ 0 & \text{si el cliente } i \text{ no es asignado al ciclo } k. \end{cases}$$

3. Problema generalizado de asignación

$$\text{Min} \sum_{k=1}^{K} \sum_{i=1}^{n} d_{ik}\, y_{ik}$$

sujeta a

$$\sum_{k=1}^{K} y_{ik} = 1, \qquad \text{para } i = 1, \ldots, n,$$
$$\sum_{i=1}^{n} a_i\, y_{ik} \le b_k, \qquad \text{para } k = 1, \ldots, K,$$
$$y_{ik} = 0 \text{ o } 1, \quad \text{para toda } i \text{ y } k,$$

donde K es la cantidad total de ciclos en la región y n es la cantidad de clientes.

Se informó que al implementar este modelo, Du Pont ahorró más de 200 millones de dólares. Un modelo matemático más complejo para resolver un problema parecido para Air Products Corporation fue descrito por Bell y otros (1983). Este estudio ganó el premio del Instituto de Práctica de las Ciencias Administrativas en Estados Unidos, en 1983.

La programación lineal (o más generalmente, la programación matemática) ha sido una herramienta importante para planeación logística en una gran variedad de problemas de administración de operaciones. A medida que la tecnología informática mejora y la computación personal, poco costosa, es cada vez más accesible para los profesionales, las técnicas de solución de propósito general, como la programación lineal, debe desempeñar un papel importante en la planeación de la producción y de las operaciones en la próxima década.

[5] En esta sección se interpreta a la programación lineal en un sentido amplio, incluyendo la programación lineal entera.

Bibliografía

Bell, W. J., L. M. Dalberto, M. L. Fisher, A. J. Greenfield, R. Jaikumar, P. Kedia, R. G. Mack y P. J. Prvtzman. "Improving the Distribution of Industrial Gases with an On-Line Computerized Routing and Scheduling Optimizer". *Interfaces* 13 (1983), pp. 4-23.

Fisher, M., A. J. Greenfield, R. Jaikumar y J. T. Uster III. "A Computerized Vehicle Routing Application". *Interfaces* 12 (1982), pp. 42-52.

Hadley, G. *Linear Programming*. Reading, MA: Addison-Wesley, 1962.

Hillier, F. S. y G. J. Lieberman. *Introduction to Operations Research*. Quinta edición. San Francisco: Holden Day, 1990.

Ragsdale, C. T. *Spreadsheet Modeling and Decision Analysis*. Segunda edición. Cincinnati: South-Western College Publishing, 1998.

Capítulo **Cuatro**

Control de inventarios sujeto al conocimiento de la demanda

Panorama general del capítulo

Propósito

Considerar métodos para controlar los inventarios de artículos individuales cuando se asume que la demanda seguirá un patrón conocido (esto es, el error de pronóstico de la demanda equivale a cero).

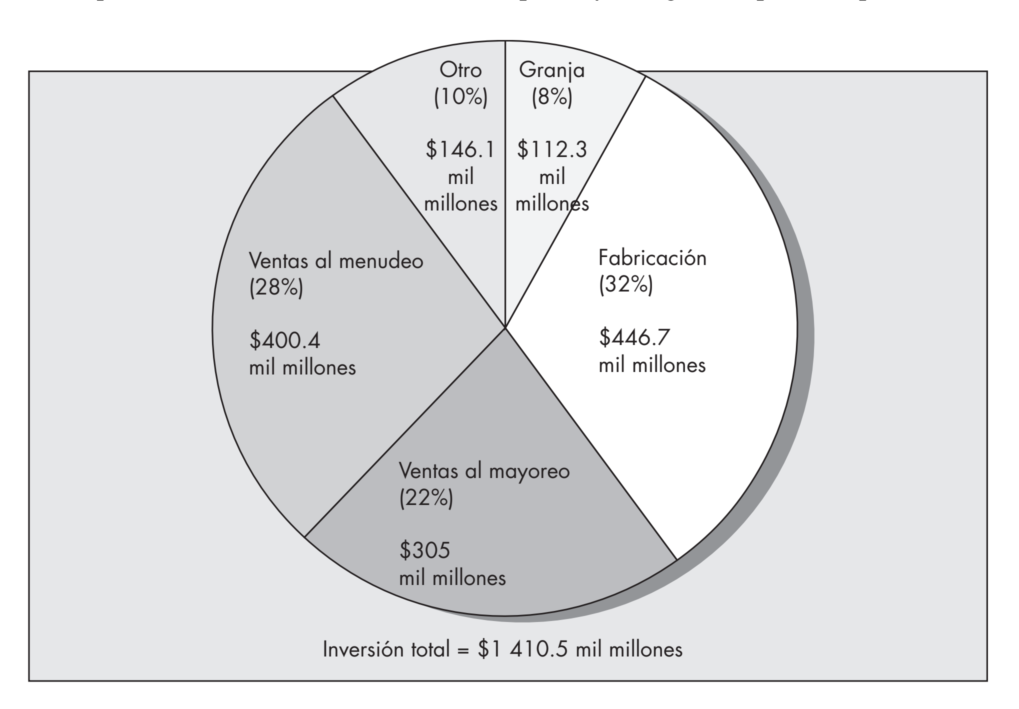
Puntos clave

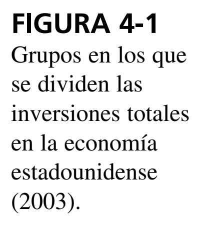
1. *Clasificación de inventarios*
   - *Materias primas.* Son los recursos necesarios para la producción o procesamiento.
   - *Componentes.* Pueden ser materias primas o subgrupos que se incluirán posteriormente en un producto final.
   - *Trabajo en proceso (WIP).* Son los inventarios que se encuentran en la planta esperando ser procesados.
   - *Artículos terminados.* Son artículos que han terminado su proceso de producción y esperan ser embarcados.
2. *¿Por qué mantener inventario?*
   - *Economías de escala.* Pedir o producir lotes grandes probablemente sea más barato que hacerlo en pequeños lotes.
   - *Incertidumbre.* Las incertidumbres en la demanda, en el tiempo de demora y en el abastecimiento. Todas ellas nos obligan a conservar inventarios.
   - *Especulación.* Los inventarios pueden acumularse anticipándonos a una elevación en su valor o costo.
   - *Transporte.* Dirigir inventarios que están en tránsito de un lugar a otro.
   - *Suavizamiento.* Como se comentó en el capítulo 3, los inventarios suavizan un patrón de demanda irregular.
   - *Logística.* Restricciones del sistema que pueden requerir mantener inventarios.
   - *Costos de control.* Mantener un inventario puede disminuir los costos que se necesitan para monitorear un sistema. (Por ejemplo, puede ser más barato hacer pedidos anuales y guardar las unidades, que realizar pedidos semanales y monitorear de cerca los pedidos y las entregas.)

3. *Características de los sistemas de inventario*
   - *Patrones de demanda.* Los dos patrones son: *a*) constante contra variable y *b*) conocido contra incierto.
   - *Reabastecimiento de tiempos de demora.* Es el tiempo entre un pedido (o la iniciación de la producción) hasta que el pedido llega (o se ha completado el proceso de producción).
   - *Tiempos de revisión.* Los puntos en el tiempo en que se revisan los niveles de inventario actuales.
   - *Tratamiento de demanda en exceso.* Cuando la demanda excede el abastecimiento; la demanda en exceso puede acumularse como pedidos atrasados o perderse.
4. *Costos relevantes*
   - *Costos de mantener el inventario.* Incluyen el costo de oportunidad de ingresos perdidos o de inversiones alternas; costos de almacenamiento físico; seguro; rotura y robo, y obsolescencia.
   - *Costos de pedido.* Comprenden dos componentes: uno fijo y uno variable. Se incurre en el componente fijo siempre que se hace un pedido positivo (o se inicia una corrida de producción); el componente variable es el costo que se paga por cada unidad, pedida o producida.
   - *Costos de penalidad.* Se incurre en ellos cuando la demanda excede el abastecimiento. En este caso, la demanda en exceso puede quedar en espera (para surtirse más adelante) o perderse. La demanda perdida da como resultado desperdiciar una utilidad, y los pedidos pendientes requieren de un registro. En ambos casos, arriesgamos la buena voluntad del cliente.
5. *El modelo EOQ básico.* El modelo EOQ data de 1915 y constituye la base de todos los modelos de control de inventarios desarrollados posteriormente. Trata del intercambio básico entre el costo fijo de realizar un pedido y el costo variable de mantener el inventario. Si h representa el costo de mantener el inventario por unidad y K el costo fijo de su disposición, entonces la cantidad de pedido que minimiza los costos por unidad es $Q = \sqrt{2K\lambda/h}$, donde λ es la tasa de demanda. Esta fórmula es muy robusta por varias razones: *a*) es una aproximación muy exacta de la cantidad óptima de pedido cuando la demanda resulta incierta (como podrá verse en el capítulo 5), y *b*) mostramos que las desviaciones de la Q óptima con frecuencia dan como resultado modestos errores de costo. Por ejemplo, un error de 25% en Q da como resultado un error de costo promedio anual por establecimiento y por mantener inventarios de sólo 2.5%.
6. *La tasa de producción finita con EOQ.* Ésta es una extensión del modelo EOQ básico que debe tomarse en cuenta cuando los artículos se producen internamente, más que cuando se piden a un proveedor externo; la tasa de producción es finita más que infinita, como se requerirá en el modelo EOQ sencillo. Mostramos que el tamaño óptimo de una corrida de producción hoy en día sigue la fórmula $Q = \sqrt{2K\lambda/h'}$ donde $h' = h(1 - \lambda/P)$ y P es la tasa de producción ($P > \lambda$). Observe que, como $h' < h$, el tamaño de lote cuando se toma en cuenta la tasa de producción excede el obtenido mediante la fórmula EOQ.
7. *Descuentos de cantidad.* Consideramos dos tipos de descuento de cantidad: todas las unidades y descuentos incrementales. En cuanto a descuentos de todas las unidades, el descuento se aplica a todas las unidades en el pedido, mientras que, en el caso de los descuentos incrementales, el descuento se aplica sólo a las unidades que están por encima del punto de equilibrio. El caso de todas las unidades con mucho es el más común en la práctica, pero es también el que más encuentra descuentos incrementales en la industria. En el caso de los descuentos de todas las unidades, el procedimiento de optimización implica buscar el punto más bajo en una curva de costos anual rota. En lo tocante a los descuentos incrementales, la curva de costo anual es continua, pero tiene derivados discontinuos.

8. *Sistemas de producto múltiple restringido de recursos.* Consideremos una tienda minorista que pide diferentes artículos pero no puede exceder un presupuesto ya fijo. Si optimizamos la cantidad del pedido de cada artículo por separado, entonces cada uno debe solicitarse de acuerdo con su valor EOQ. Sin embargo, supongamos que al hacer esto se excede el presupuesto. En esta sección se ha desarrollado un modelo que toma en cuenta explícitamente la restricción de presupuesto y ajusta al mismo los valores de EOQ. En la mayoría de los casos, la solución óptima sujeta a la restricción de presupuesto requiere una búsqueda repetitiva del multiplicador de Lagrange. Sin embargo, cuando se cumple la condición $c_1/h_1 = c_2/h_2 = \cdots = c_n/h_n$, las cantidades óptimas de un pedido son un escalamiento simple de los valores EOQ óptimos. Observe que este problema es matemáticamente idéntico a aquel en el que la restricción está en el espacio disponible, más que en el posible presupuesto.
9. *Modelos EOQ para la planeación de producción.* Supongamos que n productos distintos se producen en una sola línea de producción o máquina. Asumamos que conocemos los costos de mantener el inventario, las tasas de la demanda y las tasas de producción de cada uno de los artículos. La meta es determinar la secuencia óptima para producir los artículos, y el tamaño de lote óptimo por cada uno de los artículos para cubrir la demanda y minimizar los costos. Observe que el simple hecho de establecer un tamaño de lote para cada artículo equivalente a su valor EPQ (esto es, el tamaño óptimo de lote con una tasa de producción finita), probablemente sea subóptimo, ya que puede dar como resultado situaciones de desabasto. El problema se maneja considerando el tiempo de ciclo óptimo, T, donde asumimos que producimos exactamente un lote por cada artículo en cada ciclo. El tamaño óptimo de la corrida de producción para el artículo j es simplemente $Q_j = \lambda_j T$, donde T es el tiempo de ciclo óptimo. No obstante, el hallazgo no es cosa trivial.

La inversión actual en inventarios en Estados Unidos es enorme. En el tercer trimestre de 2003, la inversión total en dólares se estimaba en $1 410 miles de millones de dólares. La figura 4-1 muestra la inversión en inventarios desglosada por sectores de la economía. Los modelos de inventarios que estaremos comentando en este capítulo y el siguiente pueden aplicarse a todos




FIGURA 4-1
Grupos en los que se dividen las inversiones totales en la economía estadounidense (2003).

[1] *Encuesta realizada por Current Business* (octubre de 2003).

los sectores de la economía que aparecen en la figura 4-1, pero son mayormente aplicables a los sectores de manufactura, venta de mayoreo y venta de menudeo, que componen 82% del total. La inversión del billón de dólares en inventario representa 20 a 25% del producto nacional bruto anual total. Es claro que hay un potencial enorme para mejorar la eficiencia de la economía si se controlan los inventarios de manera inteligente. Las empresas que poseen métodos científicos de control de inventarios tienen una ventaja competitiva apreciable en el mercado.

Una gran parte de este libro se dedica a presentar y analizar diversos modelos matemáticos que pueden ayudar a controlar el reabastecimiento de los inventarios. En los capítulos 4 y 5 se supone que la demanda del artículo producido es externa al sistema. En la mayoría de los casos esto significa que se adquiere o produce el inventario para cumplir con las necesidades de los clientes. Sin embargo, en un ambiente de manufactura, las demandas de ciertas partes son resultado de programas de producción para ensambles de niveles superiores; las decisiones de tamaño de lote de producción en un nivel del sistema son resultado de los patrones de demanda en otros niveles. La interacción de componentes, subgrupos y productos finales desempeña un papel importante al determinar la demanda en el futuro. A este tipo de sistemas se les llama *sistemas de planeación de requerimientos de materiales* (sistemas MRP, del inglés *materials requirements planning)* o sistemas dependientes de la demanda, los cuales se describirán con detalle en el capítulo 7.

El problema fundamental de la administración de inventarios se puede describir en forma sucinta con dos preguntas: 1) ¿cuándo se debe hacer un pedido?, y 2) ¿cuánto se debe pedir? La complejidad del modelo resultante depende de los supuestos que se hagan acerca de los diversos parámetros del sistema. La mayor diferencia se da entre los modelos que suponen una demanda conocida (en este capítulo) y los que suponen demanda aleatoria (capítulo 5), aunque, como se verá, la forma de las funciones de costo y los supuestos que se hacen acerca de las características físicas del sistema también desempeñan un papel importante al determinar la complejidad del modelo resultante.

En general, los modelos que se describirán pueden usarse de manera indistinta para plantear el reabastecimiento con un proveedor externo o con producción interna. Esto significa que desde el punto de vista del modelo, el control de inventarios y la planeación de producción con frecuencia son sinónimos. Por ejemplo, los métodos de dimensionamiento de lote que se describen en el capítulo 7 también podrían haberse incluido en este capítulo. El asunto no es la etiqueta que se le ponga a una técnica, sino si dicha técnica se aplica en forma correcta al problema que se esté manejando.

4.1 TIPOS DE INVENTARIOS

Cuando se hace alusión a los inventarios en el contexto de la manufactura y la distribución, hay un esquema natural de clasificación que sugiere el valor agregado por la manufactura o el procesamiento. (Desde luego que este no es el único medio de categorizar los inventarios, pero es el más natural para las aplicaciones en manufactura.)

1. *Materias primas.* Recursos que requiere la actividad de producción o procesamiento de la empresa.
2. *Componentes.* Artículos que todavía no han sido terminados en el proceso de producción. A los componentes a veces se les llama subgrupos.
3. *Trabajo en proceso.* Inventario que espera en el sistema para procesarse o ser procesados. Los inventarios de trabajo en proceso incluyen los de componentes, y también pueden incluir algunos de materias primas. El nivel del inventario de trabajo en proceso con frecuencia se usa como medida de la eficiencia de un sistema de programación de la producción. El método justo a tiempo, que se describe con detalle en el capítulo 7, tiene por objeto reducir al mínimo el trabajo en proceso.

4. *Bienes terminados.* También llamados artículos terminados o artículos finales, son los productos finales del proceso de producción. Durante la producción se agrega valor al inventario en cada nivel de la operación de manufactura, que culmina con los artículos terminados.

El nombre o etiqueta apropiado para el inventario depende del contexto. Por ejemplo, los componentes de algunas operaciones podrían ser los productos finales de otras.

4.2 MOTIVACIÓN PARA MANTENER INVENTARIOS

1. *Economías de escala.* Imaginemos una empresa que produce una línea de artículos semejantes, como por ejemplo, filtros de aire para automóviles. Cada corrida de producción con determinado tamaño de filtro implica reconfigurar la línea de producción y recalibrar las máquinas. Como la empresa debe invertir bastante tiempo y dinero en preparar la producción de cada tamaño de filtro, se deben producir los filtros suficientes en cada corrida para justificar este costo. Ello significa que podría economizarse produciendo una cantidad relativamente elevada de artículos en cada corrida y almacenándolos para uso futuro. Lo anterior permitiría a la empresa amortizar los costos fijos de preparación, repartidos en una cantidad mayor de unidades.[2]

2. *Incertidumbres.* Con frecuencia desempeñan un papel principal en la motivación de una firma para almacenar inventarios. La incertidumbre en la demanda externa es la más importante. Por ejemplo, un detallista almacena diversos artículos para poder responder a las preferencias de los clientes. Si un cliente pide un artículo que no está disponible de inmediato, probablemente irá a cualquier otra parte. Lo peor es que quizá ese cliente jamás regrese. Al tener inventario se cuenta con un "colchón" para absorber la incertidumbre de la demanda.

Hay otras incertidumbres que también proporcionan motivos para mantener inventarios. Una es la del tiempo de demora. El tiempo de demora se define como el intervalo comprendido entre colocar un pedido y recibir lo que se solicitó. En el contexto de planeación de producción, el tiempo de demora se interpreta como el requerido para producir el artículo. Aun cuando se pueda predecir con exactitud la demanda en el futuro, la empresa necesita mantener existencias de amortiguamiento o colchón para asegurar un flujo uniforme de producción, o la continuidad de las ventas cuando los tiempos de demora de reabastecimiento sean inciertos.

Una tercera fuente importante de incertidumbre es el abastecimiento. El embargo petrolero de la OPEP a finales de los años setenta es un ejemplo del caos que puede generarse cuando se ven amenazados los canales del suministro. Dos industrias que se basaban (y continúan basándose) en el petróleo y la gasolina son las empresas eléctricas y las aerolíneas. En esas y otras industrias las empresas corrieron el riesgo de restringir sus operaciones debido a la escasez de combustible.

Otras incertidumbres que podrían hacer que una empresa almacenara inventario son la incertidumbre en la oferta de mano de obra, en el precio de los recursos y en el costo del capital.

3. *Especulación.* Si se espera que aumente el valor de un artículo o recurso natural podría ser más barato comprar grandes cantidades a los precios actuales y almacenar los artículos para usarlos después, en vez de pagar mayores precios en una fecha futura. Por ejemplo, a principios de los setenta, Westinghouse Corporation sufrió grandes pérdidas en sus contratos de construcción de centrales nucleares con varias empresas de servicio eléctrico, debido a que garantizó el suministro, a precio fijo, del uranio necesario para el funcionamiento de las plantas. Desafortunadamente para Westinghouse, el uranio se fue a las nubes en el intervalo ocurrido entre la firma de los contratos y la puesta en marcha de las plantas.

Hay otras industrias que requieren grandes cantidades de artículos costosos, mismas que han experimentado gran fluctuación en su oferta. Por ejemplo, en la producción de películas fotográficas se requiere plata. Al anticipar en forma correcta un gran aumento en el precio de

---

[2] Para este argumento se requiere que el costo de preparación sea una constante "fija". En ciertas circunstancias se puede reducir para justificar tamaños menores de lote. Esto forma la base de la filosofía justo a tiempo, que se describe con detalle en el capítulo 7.

la plata, un importante productor de películas fotográficas como lo es Kodak pudo comprar y almacenar grandes cantidades de ésta antes del aumento de los precios, y gozar de ahorros sustanciales.

El motor especulativo también puede ser un factor para una empresa que encara la posibilidad de una huelga. El costo de producción podría aumentar bastante cuando hay falta de mano de obra.

4. *Transporte.* Los inventarios en tránsito, o *proyectados*, existen porque los tiempos de transporte son positivos. Si éstos son largos, como cuando se transporta petróleo del Medio Oriente a Estados Unidos, la inversión en inventarios en tránsito puede ser apreciable. Una de las desventajas de producir en el extranjero es el mayor tiempo de transporte y, en consecuencia, el aumento en los inventarios en tránsito. Este factor ha sido determinante para que algunas compañías establezcan operaciones locales de producción.

5. *Suavizamiento.* Los cambios en el patrón de la demanda de un producto pueden ser deterministas o aleatorios. La estacionalidad es un ejemplo de variación determinista, y los cambios imprevistos de condiciones económicas pueden originar variación aleatoria. Producir y almacenar inventarios previendo un pico en la demanda puede ayudar a corregir las interrupciones causadas al cambiar las tasas de producción y los niveles de fuerza de trabajo. En el capítulo 3 se describieron, en los modelos de planeación agregada, los costos de suavizamiento y la planeación para oscilaciones previstas en la demanda.

6. *Logística.* Este término se usa para describir las razones de almacenar inventarios (que no sean las que acaban de mencionarse). Hay ciertas restricciones que pueden surgir en las compras, producción o distribución de los artículos que obligan al sistema a mantener inventarios. Uno de estos casos es el de un artículo que debe comprarse en cantidades mínimas. Otro es la logística de la manufactura; es virtualmente imposible reducir a cero todos los inventarios y esperar que haya continuidad en un proceso de manufactura.

7. *Costos de control.* Un tema importante, y que con frecuencia se olvida, es el costo de mantener el sistema de control de inventarios. Un sistema en el que se conserva más inventario no requiere el mismo nivel de control que otro en el que los niveles de inventario se mantienen al mínimo posible. Para la empresa puede resultar menos costoso, a la larga, mantener grandes inventarios de artículos poco costosos que gastar tiempo de mano de obra para mantener registros detallados de esos artículos. Aun cuando los costos de control pudieran ser un factor importante para determinar lo adecuado de determinada técnica o sistema, casi nunca se cuentan entre los modelos de inventario que se describirán.

4.3 CARACTERÍSTICAS DE LOS SISTEMAS DE INVENTARIO

1. *Demanda.* Los supuestos que se hacen acerca del comportamiento y las características de la demanda pueden ser, con frecuencia, lo más importante para determinar la complejidad del modelo de control resultante.

a) *Constante o variable.* En los modelos más sencillos de inventario se supone que la tasa de la demanda es una constante. El modelo de cantidad económica de pedido (EOQ, *economic order quantity*) y sus extensiones se basan en este supuesto. La demanda variable se presenta en distintos contextos, que incluyen la planeación agregada (capítulo 3) y la planeación de requerimientos de materiales (capítulo 7).

b) *Conocida o desconocida.* Es posible que la demanda esperada sea constante, pero al mismo tiempo aleatoria. Incierta y estocástica son sinónimos de aleatoria. Casi todos los modelos de demanda estocástica suponen que la tasa promedio de la demanda es constante. Los modelos de demanda aleatoria son, por lo general, más realistas y complejos que sus contrapartes deterministas.

2. *Tiempo de demora.* Si se piden los artículos del exterior, el tiempo de demora se define como el intervalo de tiempo que pasa desde el instante en que se hace un pedido hasta que

llega el artículo. Sin embargo, si los artículos se producen internamente, el tiempo de demora se interpreta como el tiempo requerido para producir un lote de artículos. Se usará la letra griega τ para representar el tiempo que ello demore, el cual se mide en las mismas unidades de tiempo que la demanda. Esto es, si la demanda se expresa en unidades por año, el tiempo de demora se debe expresar en años.

3. *Tiempo de revisión.* En algunos sistemas se conoce siempre el nivel actual de inventarios. Es un supuesto exacto cuando se registran las transacciones de la demanda en el momento en que ocurren. Por ejemplo, un sistema en el que los niveles de inventario siempre se conocen es el de un supermercado moderno, con lectores digitales en las cajas, que se enlaza con la base de datos de inventario a nivel almacén. A medida que un artículo se pasa por el lector, se registra la transacción en la base de datos y el nivel de inventario decrece en una unidad. A este caso lo llamaremos de *revisión continua.* En el otro caso, llamado *revisión periódica*, los niveles de inventario sólo se conocen vía puntos discretos del tiempo. Un ejemplo de revisión periódica es una tienda pequeña de víveres en la que se requiere levantamiento físico de existencias para calcular los niveles actuales de inventario disponible.

4. *Exceso de demanda.* Otra característica distintiva e importante es la manera en que reacciona el sistema hacia los excesos de demanda (esto es, demanda que no puede ser satisfecha de inmediato según los inventarios). Los supuestos más comunes son que el exceso de demanda se corre y acumula (se mantiene para ser satisfecho en el futuro) o se pierde (por lo general se satisface fuera del sistema). Entre otras posibilidades se incluye acumulación parcial (parte de la demanda se acumula y parte se pierde) o impaciencia del cliente (si el pedido del cliente no se satisface dentro de un tiempo fijo, es cancelado). La gran mayoría de los modelos de inventario, en especial los que se aplican en la práctica, suponen acumulación total del exceso de la demanda.

5. *Inventario cambiante.* En algunos casos el inventario sufre cambios a través del tiempo, que pueden afectar su utilidad. Algunos artículos tienen una vida limitada en almacén, como por ejemplo los alimentos, y otros pueden hacerse obsoletos, como las refacciones automotrices. Los modelos matemáticos que incorporan los efectos perecederos o de obsolescencia son, por lo general, muy complicados y salen del propósito de este libro. En la sección 5.8 se mencionan algunas referencias sugeridas para el lector interesado en estos modelos.

4.4 COSTOS RELEVANTES

Como nos interesa optimizar el sistema de inventarios, debemos determinar un criterio de optimización o de eficiencia adecuado. Casi todos los modelos de inventario usan la minimización del costo como criterio de optimización. Un criterio alternativo de eficiencia podría ser la maximización de la ganancia. Sin embargo, la minimización del costo y la maximización de la ganancia son criterios equivalentes en lo esencial, para la mayor parte de los problemas de control de inventario. Aunque distintos sistemas tienen características diferentes, casi todos los costos de inventario pueden colocarse en una de las siguientes tres categorías: costo de mantener inventario, costo de pedido o costo de penalización. A continuación se describe cada uno de éstos.

Costo de mantener el inventario

También se le llama costo de almacén o costo de inventario, y es la suma de todos los costos proporcionales a la cantidad de inventario disponible físicamente en cualquier punto en el tiempo. Entre los componentes del costo de mantener el inventario se incluyen diversos aspectos que en apariencia no se relacionan, entre los cuales estarían:

- El costo de suministrar el espacio físico para almacenar los artículos.
- Impuestos y seguros.
- Roturas, estropicios, deterioros y obsolescencia.
- Costo de oportunidad de una inversión alternativa.

Sucede que el último punto es el más importante para calcular los costos de mantener el inventario en la mayoría de las aplicaciones. En cierto sentido, inventario equivale a dinero en efectivo. Se debe invertir capital para comprar o producir inventario, y los niveles decrecientes de inventario dan como resultado mayor capital. Dicho capital podría invertirlo la empresa internamente, en su propio funcionamiento, o de manera externa.

¿Cuál es la tasa de interés que pudiera ganar ese capital? Todos podemos colocar nuestro dinero en una sencilla cuenta de ahorro, con 2% de tasa de interés, o posiblemente en un certificado de depósito a largo plazo, que puede redituar 5%. Podríamos ganar más invirtiendo en bonos o comprando acciones industriales a corto plazo, o traspasos de certificados de participación.

Sin embargo, generalmente la mayoría de las empresas deben ganar mayores tasas sobre sus inversiones que los individuos para ser rentables. El valor de la tasa de interés que corresponde al costo de oportunidad de la inversión alternativa se relaciona con (sin ser la misma que) varias medidas contables usuales, que incluyen la tasa interna de rendimiento, el rendimiento de los activos y la tasa límite (tasa mínima que haría atractiva una inversión para la empresa). Es muy difícil calcular la tasa de interés para el costo de oportunidad. Su valor lo estima el departamento de contabilidad y por lo general es una amalgama de las medidas contables que se mencionaron. Por comodidad usaremos el término *costo de capital* para indicar este componente del costo de mantener el inventario. Podemos imaginar que el costo de mantener el inventario es una tasa agregada de interés formada por los cuatro elementos. Por ejemplo,

28% = Costo de capital
2% = Impuestos y seguros
6% = Costo de almacenamiento
1% = Costo de roturas y deterioros
37% = Cargo total por interés

Esto se interpretaría como sigue: evaluamos un cargo de 37 centavos por cada dólar que hemos invertido en inventarios durante un periodo de un año. Sin embargo, como por lo general el inventario se mide en unidades y no en dólares, conviene expresar el costo de mantenerlo en dólares por unidad por año más que en dólares por dólar por año. Si c es el valor monetario de una unidad de inventario, I la tasa anual de interés y h el costo de mantener el inventario en dólares por unidad por año, tenemos la ecuación

$$h = Ic.$$

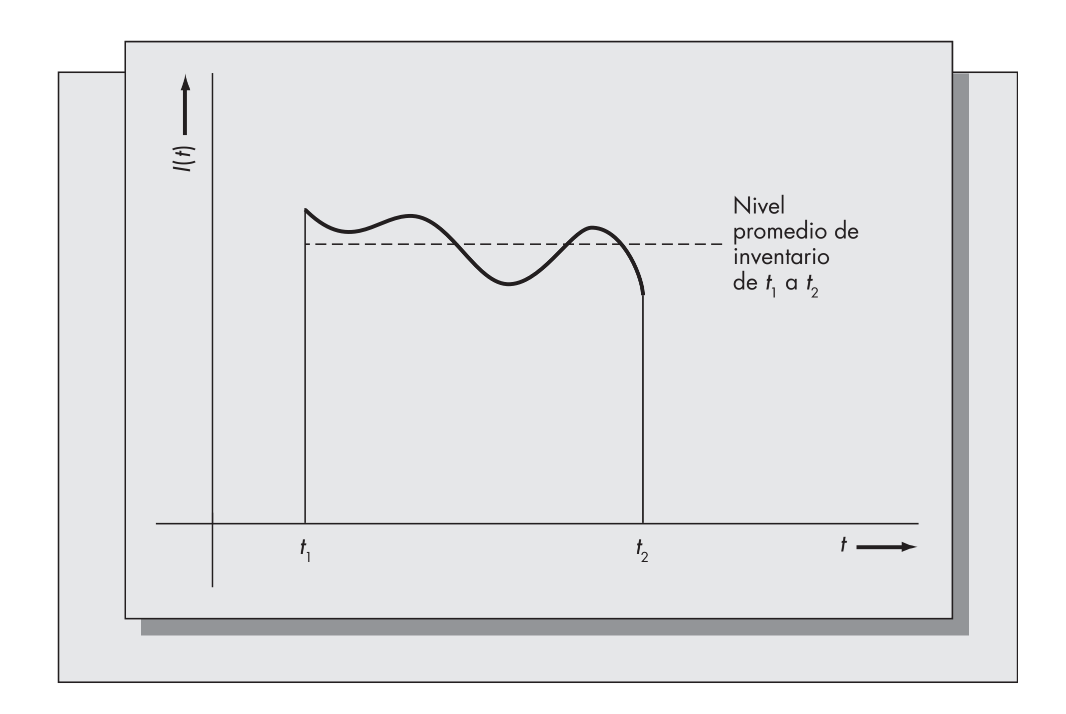
Por consiguiente, en el ejemplo anterior, un artículo valuado en 180 dólares tendría un costo anual de inventario $h = (.37)(180 \text{ dólares}) = 66.60$ dólares. Si se mantuvieran 300 artículos de esos durante cinco años, el costo total de inventario durante los cinco años sería

$$(5)(300)(66.60) = \$99\,900.$$

Este ejemplo genera una pregunta interesante. Supongamos que durante el periodo de cinco años el nivel de inventario no permaneció fijo en 300, sino que varió continuamente. Cabría esperar que los niveles de inventario cambien a través del tiempo. Decrecen cuando los artículos se usan para satisfacer la demanda, y aumentan cuando se producen unidades o llegan pedidos. ¿Cómo se calcularía el costo de mantener el inventario en ese caso? En particular, suponga que el nivel de inventario $I(t)$ durante cierto intervalo (t_1, t_2) se comporta como en la figura 4-2.

El costo de mantener el inventario incurrido en cualquier punto en el tiempo es proporcional al nivel de inventario en ese momento. En general, el costo total de mantener el inventa-

FIGURA 4-2
El inventario como una función de tiempo.



rio incurrido entre un tiempo t_1 y otro t_2 es h multiplicado por el área bajo la curva de $I(t)$. El nivel *promedio* de inventario durante el periodo (t_1, t_2) es el área bajo la curva, dividida entre $t_2 - t_1$. Para los casos que consideraremos en este capítulo se puede recurrir a la geometría sencilla para determinar esa área. Cuando $I(t)$ se describe con una recta, su valor promedio es obvio. En casos como el ilustrado en la figura 4-2, cuando la curva de $I(t)$ es compleja, el nivel promedio de inventario se determinaría calculando la integral de $I(t)$ sobre el intervalo (t_1, t_2) y dividiéndola entre $t_2 - t_1$.

Costo de pedido

El costo de mantener inventario incluye todos aquellos costos que son proporcionales a la cantidad de inventario a la mano, mientras que el costo de pedido depende de la cantidad de inventario que se pide o se produce.

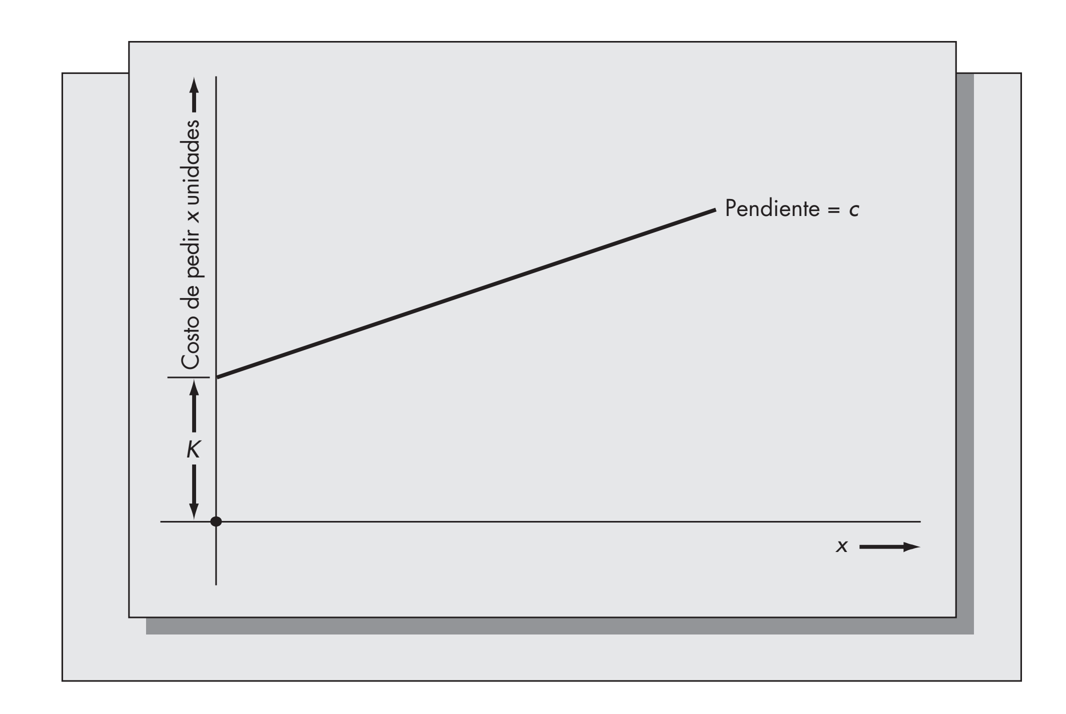
En la mayoría de las aplicaciones, el costo de pedido tiene dos componentes: uno fijo y uno variable. El costo fijo, K, es el independiente del tamaño del pedido, mientras no sea cero. El costo variable, c, es el incurrido con base en las unidades. También llamaremos a K costo de preparación y a c costo proporcional de pedido. Definiremos a $C(x)$ como el costo de pedir (o producir) x unidades. En consecuencia,

$$C(x) = \begin{cases} 0 & \text{si } x = 0, \\ K + cx & \text{si } x > 0. \end{cases}$$

La función del costo de pedido se representa en la figura 4-3.

Cuando se estima el costo de preparación se deben incluir *únicamente* los costos que sean relevantes al hecho de pedir. Por ejemplo, el costo de mantener el departamento de compras de la empresa *no* es relevante para las decisiones diarias de colocar pedidos, y no debe considerarse al estimar el costo de preparación. Ese costo del departamento de compras es un costo administrativo independiente de la decisión de colocar o no un pedido. Los costos correctos que comprenden K deben ser los gastos de contabilidad relacionados con el pedido; los costos fijos independientes del tamaño del pedido que podría requerir el vendedor; los costos de generación y recepción del pedido, y los costos de su manejo.

FIGURA 4-3
Función del costo de pedido.



Costos de penalización

El **costo de penalización**, también conocido como costo de desabasto, es el costo de carecer de suficiente inventario a mano para satisfacer una demanda *cuando se presenta*. Este costo tiene una interpretación diferente dependiendo de si la demanda en exceso es de pedidos en espera (pedidos que no pueden surtirse inmediatamente y que se mantienen en libros hasta que llegue la siguiente entrega) o perdidos (conocidos como ventas perdidas). En el caso del pedido en espera, el costo de penalización incluye todos los costos contables y/o por retraso en los que pueda incurrirse. En el caso de venta perdida, incluye la utilidad perdida que se hubiese obtenido por la venta. En cualquiera de los casos, puede incluir también un costo de *pérdida de buena voluntad*, que es una medida de satisfacción del cliente. En la práctica, puede resultar sumamente difícil estimar el componente de pérdida de buena voluntad del costo de penalización.

Usaremos el símbolo p para representar el costo de penalización y supondremos que se carga con base en las unidades. Esto es, cada vez que hay una demanda que no se puede satisfacer de inmediato, se incurre en un costo p independiente de lo que tarde satisfacerla. Un método alternativo de tener en cuenta los faltantes reside en cargar el costo de penalización en una base por unidad y por unidad de tiempo (como se hizo con el costo de mantener el inventario). Este método es adecuado si es importante el tiempo que permanece un pedido no surtido en los libros; por ejemplo, si un pedido acumulado da como resultado un paro en la línea de producción debido a la no disponibilidad de una parte. Los modelos que se consideran en este capítulo suponen que los costos de penalización se cargan sólo en base por unidad. En este capítulo no se considerarán los costos de penalización; se incluirán en el capítulo 5, pero no los que incorporan costo de penalización ponderado respecto al tiempo.

Presentaremos aquellos modelos de inventario que han tenido el mayor impacto entre la comunidad de usuarios. Muchas de las técnicas que se describen tanto en este capítulo como en el 5 forman la base de los sistemas comerciales de control de inventario, o sistemas internos. En la mayoría de los casos los modelos son lo bastante sencillos como para calcular manualmente las políticas óptimas de operación, pero con frecuencia son demasiado complejos como para plasmar los dilemas esenciales en la administración de inventarios.

Problemas para las secciones 4.1 a 4.4

1. ¿Cuáles son las dos preguntas a las que trata de contestar el control de inventarios?
2. Describa las penalizaciones de costo incurridas por una empresa que mantiene demasiado inventario, y por otra que mantiene un inventario reducido.
3. ABC, S. A., produce una línea de bicicletas de turismo. Específicamente, ¿cuáles son los cuatro tipos de inventarios (materias primas, componentes, trabajo en proceso y artículos terminados) que se tendrían en la producción de esos artículos?
4. I Carry renta camiones para movimiento y traslados. Cada camión le cuesta a la compañía un promedio de 8 000 dólares y el inventario de camiones varía mensualmente, dependiendo de la cantidad que están rentados. Durante los ocho primeros meses del año pasado, I Carry tuvo los siguientes inventarios de camiones disponibles:

Mes	Número de camiones	Mes	Número de camiones
Enero	26	Mayo	13
Febrero	38	Junio	9
Marzo	31	Julio	16
Abril	22	Agosto	5

I Carry usa una tasa anual de interés de 20% para representar su costo de capital. Los costos anuales de almacenamiento ascienden a 3% del valor de cada camión, y el costo del seguro por responsabilidad es 2%.

a) Calcule el costo total de manejo incurrido por transportes durante el periodo de enero a agosto. Suponga, para fines de su cálculo, que el costo de mantener el inventario, incurrido en cada mes, es proporcional al inventario disponible al final de ese mes.

b) Suponiendo que esos ocho meses sean representativos, estime el costo promedio anual de mantener el inventario de camiones.

5. Stationery Supplies quiere instalar un sistema de control de inventario en su almacén de Progreso. Ese almacén vende unos 1 400 artículos distintos de inventario, y sus ventas anuales aproximadas son de 80 000 dólares. El sistema de control de inventarios costaría 12 500 dólares en instalarse y unos 2 000 dólares anuales en consumos adicionales, tiempo y mantenimiento. Si los ahorros que significa el sistema se pueden representar como un porcentaje fijo de las ventas anuales, ¿cuál debería ser ese porcentaje para que el sistema se pagara por sí solo en cinco años o menos?
6. Para el caso de Stationery Supplies, descrito en el problema 5, haga una lista y describa todas las incertidumbres que motivarían a la empresa a mantener inventarios de sus 1 400 artículos.
7. Stationery Supplies pide gomas de plástico a una empresa de Nuremberg, Alemania. Se necesitan seis semanas para enviar las gomas de Alemania a Stationery Supplies. Esta empresa coloca un pedido permanente de 200 gomas cada seis meses (que se embarcan el 1 de enero y el 1 de julio).

a) Suponiendo que la política de pedidos de Stationery Supplies no ocasiona grandes acumulaciones de inventario, ni agotamientos a largo plazo, ¿cuál es la demanda anual de gomas?

b) Trace una gráfica del inventario en tránsito (esto es, el que se pidió, pero no se ha recibido) de las gomas durante un año. ¿Cuál es el inventario promedio en tránsito de las gomas durante el año?

c) Exprese el tiempo de demora de reabastecimiento en años, y multiplique la demanda anual que obtuvo en la parte *a*) por el tiempo de demora. ¿Qué observa usted acerca del resultado que se obtiene?

8. Los costos de penalización se pueden evaluar sólo en función de la cantidad de unidades de demanda que no se puede satisfacer, o en función de la cantidad de unidades, ponderada por la cantidad de tiempo que un pedido permanece en libros. Se tiene la siguiente historia de transacciones de oferta y demanda para una parte determinada.

Mes	Núm. de artículos recibidos	Demanda durante el mes
Enero	200	520
Febrero	175	1 640
Marzo	750	670
Abril	950	425
Mayo	500	280
Junio	2 050	550

Suponga que el inventario inicial al principio de enero es de 480 unidades.

a) Determine el inventario final para cada mes. Suponga que el exceso de demanda se acumula para el mes siguiente.

b) Suponga que cada vez que se solicita una unidad y no se puede surtir de inmediato se hace un cargo de 10 dólares una vez. Calcule el costo de agotamiento de inventario incurrido durante los seis meses: 1) si se pierde el exceso de demanda al final de cada mes, y 2) si se acumula el exceso de demanda al final de cada mes.

c) Suponga que cada agotamiento cuesta 10 dólares por unidad por mes, y que la demanda permanece insatisfecha. Si las demandas se surten con la base: primero en llegar, primero en servir, ¿cuál es el costo total de agotamiento incurrido durante los 6 meses usando este criterio de costo? (Suponga que la demanda se presenta al principio del mes, para fines de sus cálculos.) Observe que debe usted suponer que los excesos de demanda se acumulan en este caso, para que tenga sentido.

d) Describa bajo qué circunstancias el criterio de costos que se usó en *b*) podría funcionar, y las circunstancias bajo las cuales el criterio que se usó en *c*) podría ser adecuado.

9. HAL, Ltd., produce una línea de unidades de disco de alta capacidad para computadoras centrales. Las cajas de las unidades se producen en Hamilton, Ontario, y se embarcan a la planta principal en Toronto. HAL usa las cajas con una tasa bastante estable de 720 por año. Suponga que las cajas se embarcan en camiones que transportan 40 por viaje. Se estima que el costo fijo de cargar las cajas al camión y descargarlas en su destino cuesta 300 dólares para embarques de 120 o menos (es decir, tres camiones o menos). Cada viaje recorrido por un solo camión cuesta 160 dólares a la compañía, en tiempo de conductor, gasolina, aceite, seguros y daños en el camión.

a) Calcule los costos anuales de transporte, carga y descarga de las cajas con las siguientes políticas: 1) embarcar un camión por semana; 2) embarcar un camión lleno con la frecuencia que esto requiera; 3) embarcar tres camiones llenos con la frecuencia que esto requiera.

b) ¿Por qué razones podría suceder que la política en el punto *a*), cuyo costo anual es máximo, fuera la más conveniente desde un punto de vista de sistemas, que la política que tiene el costo mínimo anual?

4.5 EL MODELO DE CANTIDAD ECONÓMICA DE PEDIDO

Este modelo es el más sencillo y fundamental de todos los modelos de inventario. Por sus iniciales se le podría llamar modelo CEP. Describe el importante compromiso entre los costos fijos y los costos de mantener el inventario, y es la base para el análisis de sistemas más complicados.

El modelo básico

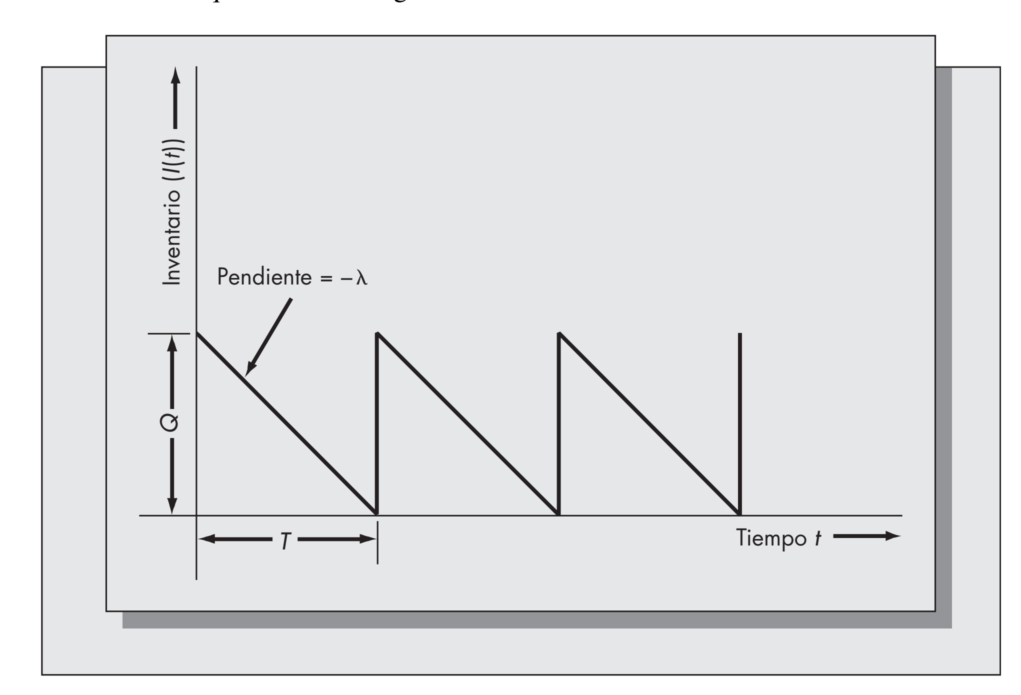
Haremos los supuestos siguientes:

1. La tasa de demanda se conoce y es una constante igual a λ unidades por unidad de tiempo. (La unidad de tiempo puede ser día, semana, mes, etc. En lo sucesivo se supondrá que la unidad de tiempo es un año, salvo que se afirme otra cosa. Sin embargo, el análisis es válido para otras unidades de tiempo, mientras se expresen todas las variables relevantes en las mismas unidades.)
2. No se permiten faltantes.
3. No hay tiempo de demora de pedido. (Este supuesto será relajado.)
4. Los costos incluyen
   a) Costo de preparación K por pedido colocado.
   b) Costo proporcional de pedido c por unidad pedida.
   c) Costo de mantener el inventario h por unidad mantenida por unidad de tiempo.

Se supondrá, sin pérdida de generalidad, que el inventario disponible en el tiempo cero es cero. No se permiten faltantes, de modo que cuando el tiempo es cero habrá que hacer un pedido. Sea Q el tamaño del pedido. En consecuencia, el inventario disponible aumentará en forma instantánea desde cero hasta Q, siendo el tiempo $t = 0$.

Veamos el siguiente tiempo cuando hay que hacer un pedido. Esta vez el inventario puede ser positivo o ser cero otra vez. Un poco de reflexión nos indica que podemos reducir los costos de mantener el inventario si esperamos hasta que el nivel de inventario llegue a cero antes de hacer un nuevo pedido. En el momento en que el inventario disponible es igual a cero, el caso parece exactamente igual que cuando $t = 0$. Si fue óptimo colocar un pedido de Q unidades en ese momento, seguirá siendo óptimo pedir Q unidades. En consecuencia, la función que describe los cambios en los niveles de inventario a través del tiempo es la forma familiar de diente de sierra que muestra la figura 4-4.

FIGURA 4-4
Niveles de inventario para el modelo de cantidad óptima de pedido.



El objetivo es elegir Q de manera que minimice el costo promedio por unidad de tiempo. A menos que se diga otra cosa, supondremos que una unidad de tiempo es un año, de modo que minimizaremos el costo anual promedio. También se pueden usar otras unidades, como día, semana o mes, siempre que todas las variables relacionadas con el tiempo se expresen en las mismas unidades. Podría pensarse que el criterio de optimización adecuado sería minimizar el costo *total* en un ciclo. Sin embargo, esto ignora el hecho de que la longitud misma del ciclo es una función de Q, y debe incluirse en forma explícita en la formulación.

A continuación se deducirá una ecuación para calcular el costo promedio anual en función del tamaño de lote Q. En cada ciclo, el costo total de pedido fijo más proporcional es $C(Q) = K + cQ$. Para obtener el costo de pedido por unidad de tiempo se divide entre la longitud del ciclo T. Como se consumen Q unidades cada ciclo a una tasa λ, en consecuencia $T = Q/\lambda$. Este resultado también puede obtenerse si se observa que la pendiente de la curva de inventario que se ilustra en la figura 4-4, $-\lambda$, es igual a la relación $-Q/T$.

Véase el costo de mantener el inventario. Como el nivel de inventario decrece linealmente desde Q hasta 0 cada ciclo, el nivel promedio de inventario durante un ciclo de pedido es $Q/2$. Como todos los ciclos son idénticos, el nivel promedio de inventario durante un horizonte de tiempo formado por muchos ciclos también es $Q/2$. Por lo tanto, el costo anual promedio, llamémosle $G(Q)$, se expresa así:

$$G(Q) = \frac{K + cQ}{T} + \frac{hQ}{2} = \frac{K + cQ}{Q/\lambda} + \frac{hQ}{2}$$

$$= \frac{K\lambda}{Q} + \lambda c + \frac{hQ}{2}.$$

Los tres términos que forman $G(Q)$ son el costo anual de preparación, el costo anual de compra y el costo anual de mantener el inventario, respectivamente.

Ahora se desea determinar la Q que minimiza a $G(Q)$. Examinando la forma de la curva $G(Q)$ se aprecia que

$$G'(Q) = -K\lambda/Q^2 + h/2$$

y

$$G''(Q) = 2K\lambda/Q^3 > 0 \qquad \text{para } Q > 0.$$

Como $G''(Q) > 0$, se sigue que $G(Q)$ es una función convexa de Q. Además, como $G'(0) = -\infty$ y $G'(\infty) = h/2$, entonces $G(Q)$ se comporta como lo ilustra la figura 4-5.

El valor óptimo de Q se presenta cuando $G'(Q) = 0$. Esto es cierto cuando $Q^2 = 2K\lambda/h$, que da como resultado

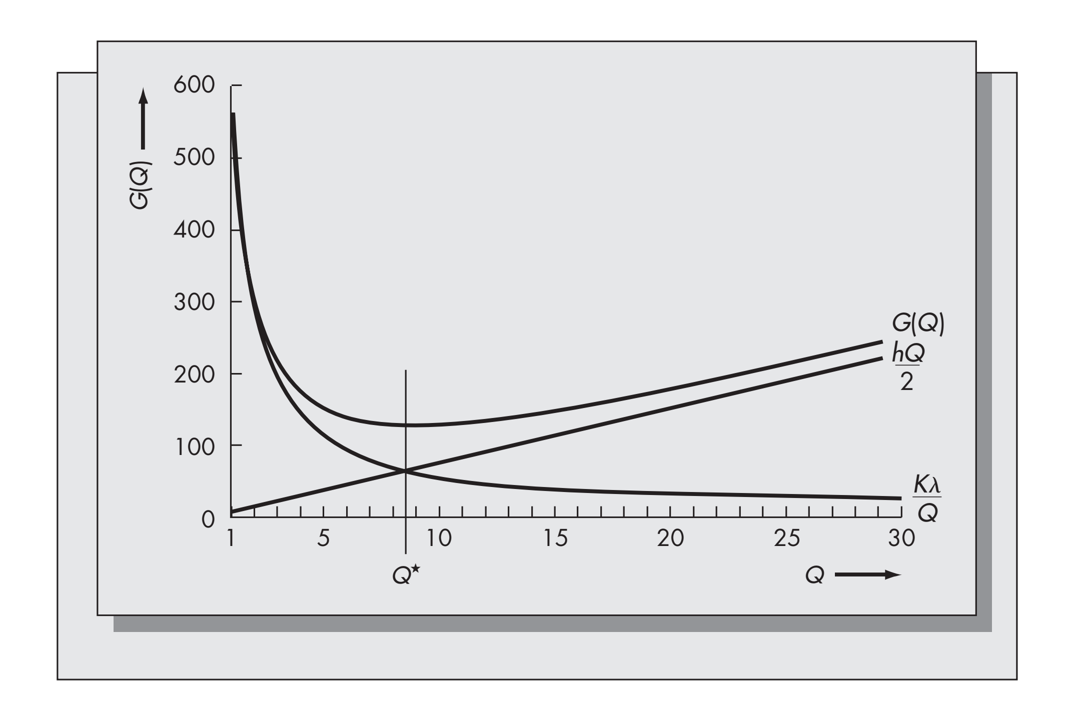
$$Q^* = \sqrt{\frac{2K\lambda}{h}}.$$

Q^* es la cantidad económica de pedido (CEP, o en inglés EOQ, *economic order quantity*). Hay varias cosas interesantes que se pueden notar:

1. En la figura 4-5, las curvas que corresponden al componente fijo del costo de pedido, $K\lambda/Q$, y el componente de costo de mantener el inventario, $hQ/2$, también están incluidas. Observe que Q^* es el valor de Q cuando se intersecan las dos curvas. (Si el lector iguala $hQ/2$ con $K\lambda/Q$ y despeja Q, obtendrá la fórmula de CEP.) En general, el mínimo de la suma de dos funciones *no* se presenta en la intersección de ellas. Es una coincidencia interesante que sí sucede en este caso.

2. Observe que el componente proporcional del costo de pedido, c, no aparece explícitamente en la ecuación de Q^*. Esto se debe a que el término λc que aparece en la definición de $G(Q)$ es independiente de Q. Como todas las políticas factibles deben reabastecer el inventario

FIGURA 4-5
La función de costo promedio anual $G(Q)$.



según la tasa de la demanda, el costo proporcional de pedido incurrido por unidad de tiempo es λc, y es independiente de Q. Como λc es una constante, por lo general no se tiene en cuenta cuando se calculan costos promedio. Observe que c *sí* afecta indirectamente al valor de Q^*, porque h aparece en la fórmula de cantidad económica de pedido y $h = Ic$.

Ejemplo *4.1*

En la papelería de una universidad se venden lápices #2 con una tasa constante de 60 por semana. A la papelería le cuestan .02 dólares cada uno y los vende a .15 dólares por pieza. Cuesta 12 dólares iniciar un pedido, y los costos de mantener inventario se basan en una tasa anual de interés de 25%. Calcule la cantidad óptima de lápices que debe comprar la papelería, así como el tiempo entre la colocación de los pedidos. ¿Cuáles son los costos anuales de mantener el inventario y de preparación para este artículo?

Solución

Primero se convierte la demanda a una tasa anual, para que sea consistente con los cargos por intereses que se hacen cada año. (También podría convertirse la tasa anual de interés a tasa semanal de interés.) La tasa anual de demanda es $\lambda = (60)(52) = 3\,120$. El costo h de mantener el inventario es el producto de la tasa de interés anual y el costo variable del artículo. Por consiguiente, $h = (0.25)(0.02) = 0.005$. Sustituyendo en la fórmula de la cantidad económica de pedido se obtiene

$$Q^* = \sqrt{\frac{2K\lambda}{h}} = \sqrt{\frac{(2)(12)(3\,120)}{0.005}} = 3\,870.$$

El tiempo del ciclo es $T = Q/\lambda = 3\,870/3\,120 = 1.24$ años. El costo anual promedio de mantener el inventario es $h(Q/2) = 0.005(3\,870/2) = \$9.675$ dólares. El costo anual promedio de preparación es $K\lambda/Q$, que también es 9.675 dólares.

El ejemplo 4.1 ilustra algunos de los problemas que pueden surgir al emplear modelos sencillos. La solución óptima indica pedir casi 4 000 lápices cada 15 meses. Aun cuando este valor de Q minimiza los costos anuales de mantener el inventario y de preparación, podría no ser factible: puede ser que la papelería no tenga espacio para almacenar 4 000 lápices. Los modelos sencillos no pueden tener en cuenta todas las restricciones presentes en un problema real. Por este motivo, se debe considerar cada solución en su contexto y, si es necesario, modificarla para ajustar su aplicación.

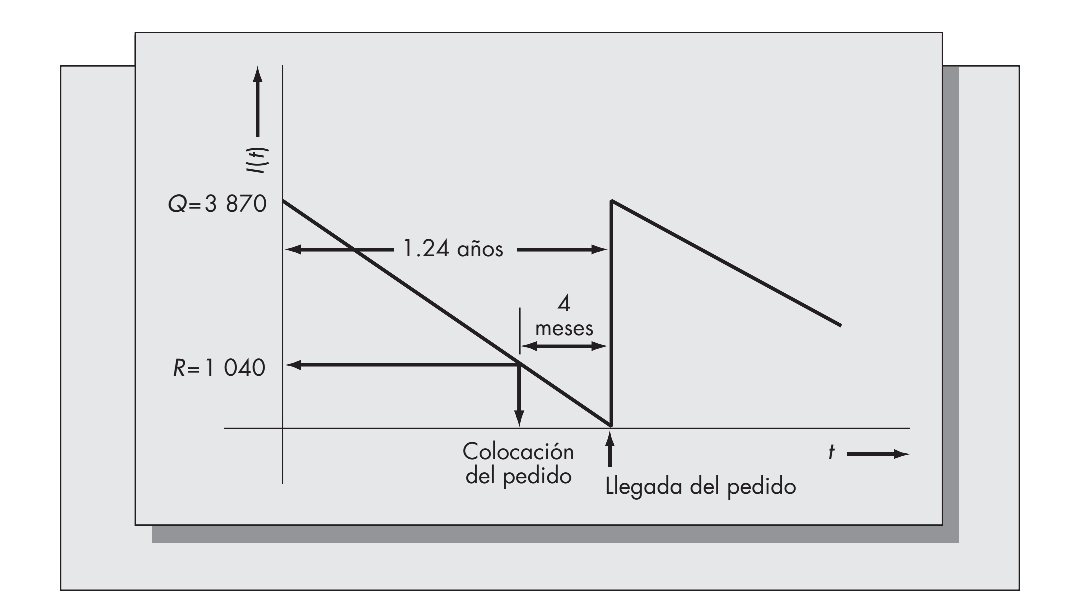


FIGURA 4-6
Cálculo de punto de reorden para el ejemplo 4.1.

Observe también que la solución óptima no dependió del precio de venta de .15 dólares. Aun cuando cada lápiz se vendiera en 2 dólares seguiría recomendándose la misma cantidad de pedido, porque se supone que los lápices se venden con una tasa de 60 por semana, independientemente de su precio. Naturalmente eso es una simplificación de la realidad. Es razonable suponer que la demanda es relativamente estable para un cierto intervalo de precios. Los modelos de inventario incorporan explícitamente el precio de venta en la formulación sólo cuando el establecimiento del precio se incluye como parte de la optimización.

Inclusión del tiempo de demora del pedido

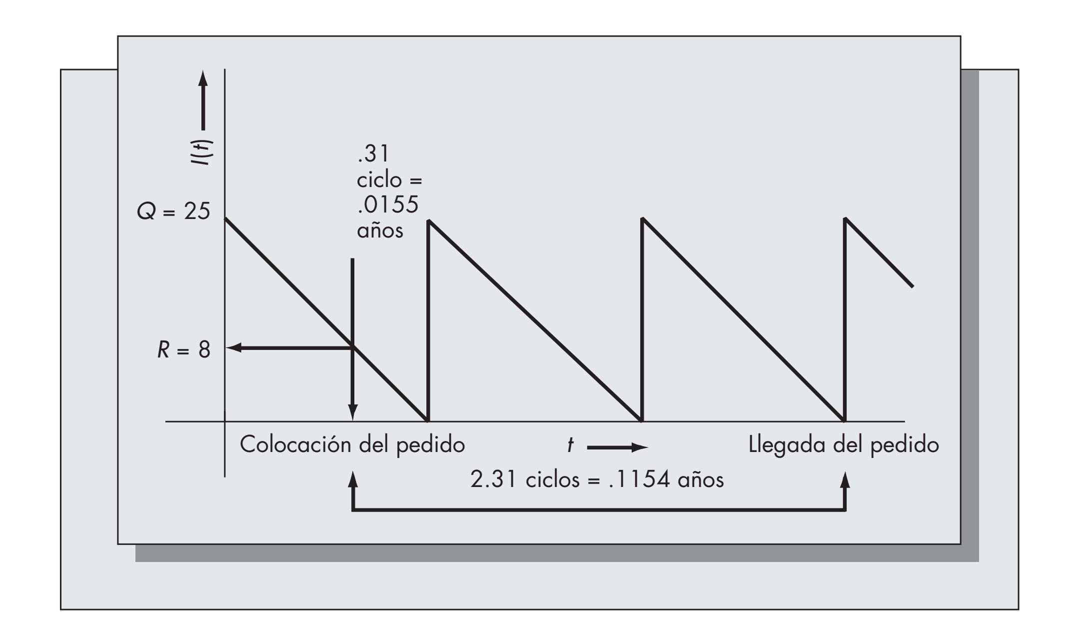
Uno de los supuestos para deducir el modelo de cantidad económica de pedido fue que no hay tiempo de demora del pedido. Ahora se rebajará ese supuesto. En el ejemplo 4.1 supondremos que hay que pedir los lápices con cuatro meses de anticipación. Si se hiciera el pedido exactamente cuatro meses antes del fin de cada ciclo, los lápices llegarían exactamente en el mismo punto en el tiempo que en el caso de tiempo de demora igual a cero. El tiempo óptimo de colocación del pedido para el ejemplo 4.1 se ilustra en la figura 4-6.

Más que estipular cuándo debe hacerse un pedido en relación con el final de un ciclo, conviene más indicar el "resurtido" o "reorden", en función del inventario disponible. Se definirá R, el punto de reorden, como el nivel del inventario disponible en el instante en que se tiene que hacer un pedido. En la figura 4-6 se aprecia que R es el producto del tiempo de la demora por la tasa de la demanda ($R = \lambda\tau$). Por ejemplo, $R = (3\ 120)(0.3333) = 1\ 040$. Obsérvese que el tiempo de demora se ha convertido a años antes de multiplicarlo. *Todas las variables relevantes siempre se expresan en las mismas unidades de tiempo.*

La determinación del punto de reorden es más difícil cuando el tiempo de demora es mayor que un ciclo. Veamos el caso de un artículo cuya cantidad económica de pedido es 25, tiene una tasa de demanda de 500 unidades por año y un tiempo de demora de seis semanas. El tiempo del ciclo es $T = 25/500 = 0.05$ año, es decir, 2.6 semanas. Si se forma la relación τ/T, se obtiene 2.31. Esto significa que hay exactamente 2.31 ciclos en el tiempo de demora. Cada pedido debe colocarse 2.31 ciclos por adelantado (véase la figura 4-7).

Observe que, para calcular el punto de reorden, resulta exactamente igual que colocar el pedido *0.31 ciclos por adelantado.* Esto sucede porque el nivel de inventario disponible es igual si estamos en el punto 2.31 o 0.31 ciclos antes de la llegada de un pedido. En este caso,

FIGURA 4-7
Cálculo del punto de reorden cuando los tiempos de demora son mayores que un ciclo.



0.31 ciclos es 0.0155 años, y así se obtiene un punto de reorden $R = (0.0155)(500) = 7.75 \approx 8$. En general, cuando $\tau > T$, se usa el siguiente procedimiento:

a) Formar la relación τ/T.

b) Tener en cuenta el residuo fraccionario de la relación. Multiplicar este residuo fraccionario por la longitud del ciclo para regresar a las unidades de años.

c) Multiplicar el resultado del paso *b*) por la tasa de demanda para obtener el punto de reorden.

Sensibilidad

En esta sección se examina cuán sensible es la función de costo anual a los errores en el cálculo de Q. Regresemos al ejemplo 4.1 y supongamos que la papelería pide los lápices en lotes de 1 000 y no de 3 870 como indica la solución óptima. ¿En qué costo adicional se está incurriendo al usar una solución subóptima? Para responder la pregunta se considerará $G(Q)$. Al sustituir $Q = 1\,000$, puede calcularse la función de costo anual promedio para este tamaño de lote y compararlo con el costo óptimo con miras a determinar la magnitud de penalización. Entonces

$$G(Q) = K\lambda/Q + hQ/2 = (12)(3\,120)/1\,000 + (0.005)(1\,000)/2 = \$39.94,$$

que es bastante mayor que el costo óptimo de 19.35 dólares.

Se puede determinar así la penalización de costo para soluciones subóptimas, para cualquier problema en particular. Sin embargo, es más instructivo y cómodo obtener una solución universal del problema de la sensibilidad. Esto se hará deduciendo una ecuación para calcular el cociente entre el costo subóptimo y el costo óptimo, en función del cociente entre las cantidades óptima y subóptima de pedido. Sea G^* el costo promedio anual de mantener inventario y de preparación en la solución óptima. Entonces

$$G^* = K\lambda/Q^* + hQ^*/2 = \frac{K\lambda}{\sqrt{2K\lambda/h}} + \frac{h}{2}\sqrt{\frac{2K\lambda}{h}} = 2\sqrt{\frac{K\lambda h}{2}}$$
$$= \sqrt{2K\lambda h}.$$

Por consiguiente, para cualquier Q,

$$\frac{G(Q)}{G^*} = \frac{K\lambda/Q + hQ/2}{\sqrt{2K\lambda h}}$$

$$= \frac{1}{2Q}\sqrt{\frac{2K\lambda}{h}} + \frac{Q}{2}\sqrt{\frac{h}{2K\lambda}}$$

$$= \frac{Q^*}{2Q} + \frac{Q}{2Q^*}$$

$$= \frac{1}{2}\left[\frac{Q^*}{Q} + \frac{Q}{Q^*}\right].$$

Para observar cómo se utiliza este resultado, véase el caso en que se usa un tamaño subóptimo de lote, en el ejemplo 4.1. La solución óptima fue $Q^* = 3\,870$, y se deseaba evaluar el error en el costo al usar $Q = 1\,000$. Al formar la relación Q^*/Q se obtiene 3.87. Por consiguiente, $G(Q)/G^* = (.5)(3.87 + 1/3.87) = (.5)(4.128) = 2.06$. Esto significa que el costo promedio anual de mantener el inventario y de preparación cuando $Q = 1\,000$ es 2.06 veces mayor que el costo óptimo anual de mantener inventario y de preparación.

En general, la función de costo $G(Q)$ es relativamente insensible a los errores en Q. Por ejemplo, si Q sube al doble de lo que debería ser, $Q/Q^* = 2$, y $G/G^* = 1.25$. En consecuencia, un error del 100% en el valor de Q da como resultado un error de sólo 25% en los costos anuales de inventario y de preparación. Observe que se obtiene el mismo resultado si Q es la mitad de Q^*, porque $Q^*/Q = 2$. Sin embargo, lo anterior *no* implica que la función del costo promedio anual sea simétrica. De hecho, suponga que la cantidad de pedido difería de la óptima en ΔQ unidades. Un valor de $Q = Q^* + \Delta Q$ debe dar como resultado un costo promedio anual *menor* que cuando $Q = Q^* - \Delta Q$.

Cantidad económica de pedido y JIT

Como resultado en buena medida del sistema kanban de Toyota, está difundiéndose una nueva filosofía acerca del papel y la importancia de los inventarios en los ambientes de manufactura. Esta filosofía, llamada *justo a tiempo* (JIT, *just-in-time*), predica que el exceso de los inventarios de trabajo en proceso no es deseable, que los inventarios deben reducirse a lo esencial. (Describimos con más detalle la filosofía justo a tiempo en el capítulo 1, y describiremos la mecánica del sistema kanban en el capítulo 7.) La cantidad económica de pedido es el resultado de la concepción tradicional acerca de inventarios y economías de escala en la teoría económica. ¿Se oponen entre sí los métodos de CEP y JIT?

Los defensores señalan que una parte esencial de la implementación del JIT es la reducción de los tiempos de preparación, y con ello los costos de este rubro. A medida que los costos de preparación decrecen, la teoría tradicional de la cantidad económica de pedido indica que deben reducirse los tamaños de lote. En este sentido, las dos maneras de pensar son compatibles. Sin embargo, en algunas ocasiones pueden no serlo. Creemos que hay un valor sustancial en el método JIT que no se puede incorporar con facilidad en un modelo matemático. Se pueden identificar y corregir problemas de calidad antes de que se acumulen inventarios de partes defectuosas. Las plantas pueden ser más flexibles si no se les carga con exceso de inventarios en proceso. Ciertamente, el éxito de Toyota con el método JIT, demostrado con los costos de inventario por automóvil, bastante más bajos que los normales en Estados Unidos, es una prueba del valor del JIT.

Sin embargo, creemos que todo nuevo método debe incorporarse con cuidado a la empresa, y no adoptarse a ciegas sin evaluar sus consecuencias y su adecuación. Aunque el justo a tiempo es un nuevo y excitante desarrollo en la administración de materiales, no siempre es el mejor método. Los principios en que se basa la cantidad económica de pedido (y MRP, la planeación de requerimientos de materiales que se describe en el capítulo 7) son sólidos y no deben pasarse por alto. El siguiente ejemplo ilustra este punto.

Ejemplo *4.2*

La planta de Rahway, Nueva Jersey, de Metalcase, fabricante de muebles de oficina, produce escritorios metálicos en un volumen de 200 mensuales. Cada escritorio requiere 40 tornillos Phillips para metal, que se compran a un proveedor de Carolina del Norte. Cuestan .03 dólares cada uno. Los costos fijos de entrega, y los de recepción y almacenamiento de los embarques de tornillos, ascienden a 100 dólares aproximadamente por embarque, y son independientes del tamaño del lote. La empresa aplica una tasa de interés de 25% para calcular sus costos de mantenimiento de inventario. A Metalcase le gustaría hacer un pedido fuerte al proveedor, y considera varias alternativas. ¿Cuál es el tamaño de pedido que deben usar?

Solución

Primero calcularemos la cantidad económica de pedido. La demanda anual de tornillos es

$$(200)(12)(40) = 96\ 000.$$

El costo anual de mantener el inventario por tornillo es (0.25)(0.03) = 0.0075. Según la fórmula para calcular la CEP, el tamaño óptimo del lote es

$$Q^* = \sqrt{\frac{(2)(100)(96\ 000)}{0.0075}} = 50\ 597.$$

Observe que el tiempo del ciclo es $T = Q/\lambda = 50\ 597/96\ 000 = 0.53$ años, más o menos una vez cada seis meses. Por consiguiente, la política óptima pide reabastecer los tornillos más o menos dos veces al año. Un método JIT sería pedir los tornillos con tanta frecuencia como fuera posible, para minimizar el inventario disponible en la planta. Al implementar ese método cabría pensar en una política de entregas mensuales. Sin embargo, esa política tiene poco sentido en este contexto, porque requeriría 52 entregas por año, e incurriría en costos anuales de preparación de 5 200 dólares. La solución CEP produce un costo anual total de preparación e inventario menor que 400 dólares. Para un artículo de bajo valor como éste, con grandes costos fijos de pedido, son inadecuados los tamaños pequeños de lote que aconseja JIT. La esencia es que no se debe adoptar a ciegas un solo método para todas las situaciones. El éxito de un método en un contexto no asegura su adecuación en todos los demás casos.

Problemas para la sección 4.5

10. Una cafetería vende café colombiano a una tasa constante de 280 libras anuales. Los granos se compran a un proveedor local a 2.40 dólares por libra. La cafetería estima que le cuesta 45 dólares colocar un pedido de café, entre papelería y horas de trabajo, y que los costos de mantener el inventario se basan en una tasa anual de interés de 20%.

    a) Calcule la cantidad óptima de pedido para el café colombiano.
    b) ¿Cuál es el tiempo entre los pedidos?
    c) ¿Cuál es el costo anual promedio de inventario y preparación debido a este artículo?
    d) Si el tiempo de demora del reabastecimiento es tres semanas, calcule el nivel de reorden, con base en el inventario disponible.

11. Para el caso del problema 10, trace una gráfica de la cantidad de inventario bajo pedido. Con ella determine la cantidad promedio de inventario bajo pedido. También calcule la demanda durante el tiempo de demora del reabastecimiento. ¿Son distintas esas dos cantidades?

12. Un gran taller automotriz instala unos 1 250 silenciadores de escape por año, 18% de los cuales son para automóviles importados. Dichos silenciadores se compran a un solo abastecedor local, a 18.50 dólares cada uno. El taller aplica un costo de inventario basado en 25% de tasa de interés anual. El costo de preparación para colocar un pedido se estima en 28 dólares.

    a) Calcule la cantidad óptima de silenciadores de escape para autos importados que debe comprar el taller cada vez que coloca un pedido, y el tiempo entre la colocación de pedidos.

b) Si el tiempo de demora para el reabastecimiento es de seis semanas, ¿cuál es el punto de reorden basado en el nivel de inventario disponible?

c) La política actual de pedidos es comprar los silenciadores para autos importados una vez al año. ¿Cuáles son los costos adicionales de inventario y preparación en que incurre esa política?

13. Regrese a la cafetería del problema 10. Suponga que su costo de preparación para los pedidos realmente es sólo 15 dólares. Calcule el error incurrido al estimar el costo anual de inventario y preparación como resultado de usar el otro valor de *K*. (Observe que esto implica que su política actual de pedidos es subóptima.)

14. Un taller local de maquinado compra tuercas hexagonales y tornillos de alta resistencia al mismo proveedor. Las tuercas cuestan 15 centavos cada una y los tornillos 38 centavos por pieza. Se estima que preparar todos los pedidos cuesta 100 dólares, incluyendo el costo de rastrear y recibir los pedidos. Los costos de mantener el inventario se basan en una tasa de interés anual de 25%. El taller usa un promedio de 20 000 tuercas hexagonales y 14 000 tornillos de alta resistencia cada año.

a) Calcule el tamaño óptimo de los pedidos de tuercas y tornillos, y el intervalo óptimo entre la colocación de pedidos para estos dos artículos.

b) Si ambos artículos se piden y reciben en forma simultánea, el costo de preparación de 100 dólares se aplica al pedido combinado. Compare el costo anual promedio de inventario y preparación, si se piden esos artículos por separado, si se piden ambos cuando se deban pedir las tuercas hexagonales y si se piden ambos cuando se deban pedir los tornillos de alta resistencia.

15. Gastronómica David recibe por avión salamis marca Hebrew National, con regularidad, para satisfacer una creciente demanda en Silicon Valley. David Gold, el dueño, estima que la demanda de salamis es bastante estable: 175 por mes. Esos salamis le cuestan 1.85 dólares cada uno. El costo fijo de llamar a su hermano en Nueva York y hacer que los envíen por avión es de 200 dólares. Se necesitan tres semanas para recibir un pedido. Irving Wu, su contador, recomienda un costo anual de capital de 22%, un costo de espacio de almacenamiento de 3% y un costo de 2% para impuestos y seguros. Estos costos se basan en el valor del artículo.

a) ¿Cuántos salamis debe recibir Gold, y con qué frecuencia debe pedirlos?

b) ¿Cuántos salamis debe tener Gold disponibles cuando telefonee para que su hermano le envíe otra remesa?

c) Suponga que los salamis se venden a 3.00 dólares cada uno. ¿Le son redituables a Gold? En caso afirmativo, ¿qué ganancia anual puede esperar con este artículo? (Suponga que opera el sistema en forma óptima.)

d) Si los salamis duran sólo cuatro semanas en almacenamiento, ¿cuál es el riesgo con la política que elaboró usted en el punto *a*)? ¿Cuál política habrá que usar en este caso? ¿Sigue siendo redituable ese artículo?

16. En vista de los resultados obtenidos en la sección sobre análisis de sensibilidad, describa la siguiente afirmación de un gerente de control de inventarios: "Si los tamaños de mis lotes van a ser erróneos, prefiero equivocarme hacia arriba que hacia abajo".

4.6 EXTENSIÓN A UNA TASA FINITA DE PRODUCCIÓN

Un supuesto implícito en el modelo sencillo de cantidad económica de pedido es que los artículos se obtienen con un abastecedor externo. Cuando así sucede, es razonable suponer que todo el lote se entrega de una vez. Sin embargo, si deseamos aplicar la fórmula de la CEP cuando

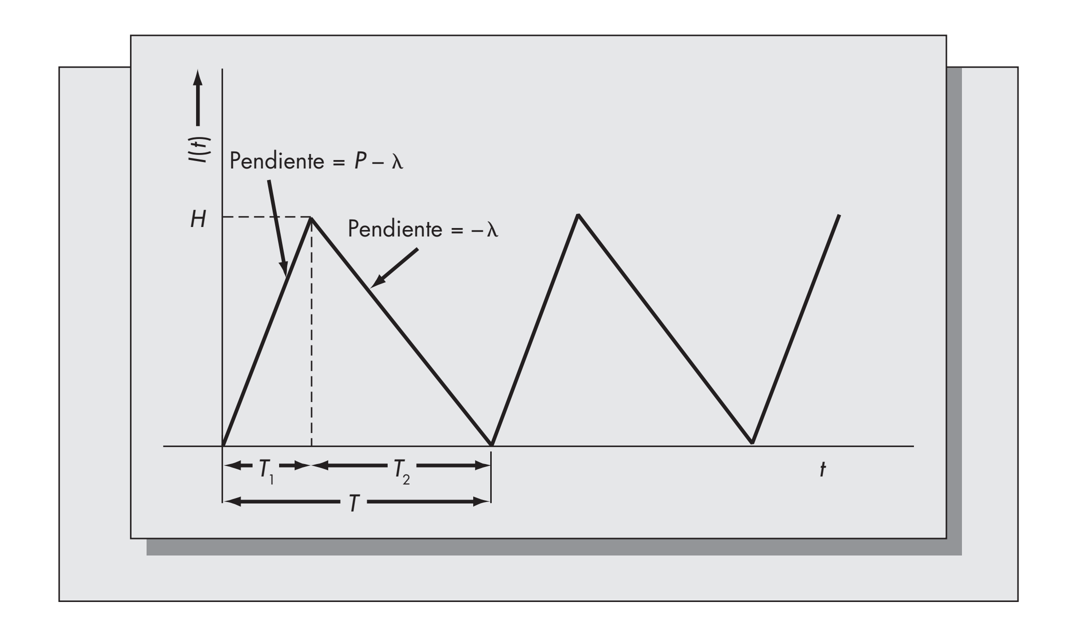


FIGURA 4-8
Niveles de inventario para el modelo de tasa finita de producción.

las unidades se producen internamente, supondríamos de hecho que la tasa de producción es infinita. Es probable que este supuesto sea satisfactorio como aproximación, cuando la tasa de producción es mucho mayor que la de demanda. Sin embargo, si la tasa de producción es comparable a la tasa de demanda, la fórmula simple de CEP producirá resultados incorrectos.

Suponga que se producen artículos a la tasa P durante una corrida de producción. Por factibilidad, se requiere que $P > \lambda$. Todos los demás supuestos serán idénticos a los de la deducción de la fórmula sencilla. Cuando las unidades se producen internamente, la forma de la curva que describe los niveles de inventario en función del tiempo es un poco distinta que la de diente de sierra de la figura 4-4. El cambio en el nivel de inventario a través del tiempo, para el caso de tasa finita de producción se ilustra en la figura 4-8.

Sea Q el tamaño de cada corrida de producción, y T la longitud del ciclo, el tiempo entre arranques sucesivos de producción. Entonces $T = T_1 + T_2$, siendo T_1 el tiempo de subida (tiempo de producción) y T_2 es el tiempo de bajada (tiempo de consumo). Observe que el nivel máximo de inventario disponible durante un ciclo *no* es Q.

La cantidad de unidades consumidas en cada ciclo es λT, que debe ser igual a la de unidades producidas cada ciclo Q. En consecuencia, $Q = \lambda T$, o $T = Q/\lambda$. Definimos H como el nivel máximo de inventario disponible. Como los artículos se producen a la tasa P durante el tiempo T_1, entonces $Q = PT_1$, es decir, $T_1 = Q/P$. En la figura 4-8 se aprecia que $H/T_1 = P - \lambda$. Esto se debe a la definición de la pendiente: distancia vertical entre distancia horizontal. Al sustituir $T_1 = Q/P$ y despejar H se obtiene $H = Q(1 - \lambda/P)$.

Ahora deduciremos una ecuación de la función del costo anual promedio. Como el nivel promedio de inventarios es $H/2$, entonces

$$G(Q) = \frac{K}{T} + \frac{hH}{2} = \frac{K\lambda}{Q} + \frac{hQ}{2}(1 - \lambda/P).$$

Obsérvese que si definimos $h' = h(1 - \lambda/P)$, entonces esa función $G(Q)$ es idéntica a la del caso de tasa infinita de producción, en el que se sustituye h' por h. En consecuencia,

$$Q^* = \sqrt{\frac{2K\lambda}{h'}}.$$

Ejemplo *4.3*

Una empresa local produce una memoria programable EPROM para varios clientes industriales. Dicha empresa ha experimentado una demanda relativamente constante de 2 500 unidades de su producto por año. La EPROM se produce a una tasa de 10 000 unidades al año. El departamento de contabilidad ha estimado que cuesta 50 dólares iniciar una corrida de producción, que la fabricación de cada unidad le cuesta 2 dólares a la empresa y que el costo de mantener el inventario se basa en una tasa anual de 30% de interés. Calcule el tamaño óptimo de una corrida de producción, su duración y el costo anual promedio de inventario y preparación. ¿Cuál es el nivel máximo de inventario disponible de memorias?

Solución

Primero calcularemos $h = (0.3)(2) = 0.6$ por unidad y por año. El costo modificado de mantener el inventario es $h' = h(1 - \lambda/P) = (0.6)(1 - 2\ 500/10\ 000) = 0.45$. Al sustituir en la fórmula de la CEP y sustituir h' en vez de h, obtenemos $Q^* = 745$. Observe que la CEP simple es igual a 645, 14% menor.

El tiempo entre las corridas de producción es $T = Q/\lambda = 745/2\ 500 = 0.298$ año. El tiempo de subida de cada ciclo es $T_1 = Q/P = 745/10\ 000 = 0.0745$ años, y el tiempo de bajada de cada ciclo es $T_2 = T - T_1 = 0.2235$ año.

El costo anual promedio de mantener y preparar el inventario es

$$G(Q^*) = \frac{K\lambda}{Q^*} + \frac{h'Q^*}{2} = \frac{(50)(2\ 500)}{745} + \frac{(0.45)(745)}{2} = 335.41.$$

El nivel máximo de inventario disponible es $H = Q^*(1 - \lambda/P) = 559$ unidades.

Problemas para la sección 4.6

17. Química Wod produce una sustancia que se usa como fertilizante para prados. Se pueden generar 10 000 libras diarias de esa sustancia. Su demanda anual es de 0.6 millones de libras por año. El costo fijo de preparar una corrida de producción del fertilizante es 1 500 dólares, y el costo variable de producción es 3.50 dólares por libra. La empresa usa una tasa de interés de 22% anual, para tener en cuenta el costo del capital, y los costos de almacenamiento y manejo de la sustancia ascienden a 12% del valor del producto. Suponga que hay 250 días hábiles por año.

    a) ¿Cuál es el tamaño óptimo de corrida de producción para esta sustancia en especial?

    b) ¿Qué proporción de cada ciclo de producción consiste en tiempo de subida, y cuál es la del tiempo de bajada?

    c) ¿Cuál es el costo promedio anual de inventario y preparación que se atribuye a este artículo? Si la empresa lo vende a 3.90 dólares por libra, ¿cuál es la ganancia anual que obtiene por este artículo?

18. Calcule el tamaño de lote que resultaría en el problema 17 si se supusiera que la tasa de producción es infinita. ¿Cuál es el costo anual promedio adicional en que se incurriría al usar este tamaño de lote y no el que se calculó en el problema 17?

19. HAL, Ltd., la empresa del problema 9, puede producir cajas para unidades de disco en la planta de Hamilton, Ontario, a una tasa de 150 cajas por mes. Producir esas cajas le cuesta 85 dólares, y el costo de preparación para iniciar una corrida de producción es 700 dólares. Suponga una tasa anual de interés de 28% para determinar el costo de inventario.

    a) ¿Cuál es la cantidad óptima de cajas que debe producir HAL en cada corrida?

    b) Calcule el tiempo entre el inicio de corridas de producción, el tiempo dedicado a la producción y el tiempo de bajada de cada ciclo de producción.

    c) ¿Cuál es la inversión máxima en efectivo (dólares) por concepto de cajas que tiene HAL en cualquier momento?

20. Filter Systems produce filtros de aire para automóviles nacionales e importados. Un filtro, número de parte JJ39877, lo consume Servicio Rayo, bajo contrato exclusivo de suministro, con la frecuencia constante de 200 unidades mensuales. Filter Systems puede producir ese filtro a una tasa de 50 por hora. El tiempo de preparación para cambiar los ajustes del equipo es 1.5 horas. El tiempo de mano de obra (que incluye gastos administrativos) se carga a 55 dólares por hora, y el tiempo inactivo de la planta durante las preparaciones se estima en 100 dólares por hora, por concepto de ganancias perdidas.

    Filter Systems ha establecido una tasa de interés anual de 22% para determinar sus costos de inventario. A la empresa le cuesta 2.50 dólares producir cada filtro, y lo vende a 5.50 dólares por pieza a Servicio Rayo. Suponga en sus cálculos días de seis horas de trabajo, 20 días hábiles por mes y 12 meses por año.

    a) ¿Cuántos filtros JJ39877 debe fabricar Filter Systems en cada corrida de producción de esta parte para minimizar los costos anuales de mantener el inventario y de preparación?

    b) Suponiendo que fabrica la cantidad óptima de filtros en cada corrida, ¿cuál es el nivel máximo de inventario disponible de esos filtros que tiene la empresa en cualquier momento?

    c) ¿Qué porcentaje del tiempo de trabajo usa esta empresa para producir estos filtros, suponiendo que se aplica la política descrita en la parte *d*?

4.7 MODELOS DE DESCUENTO POR CANTIDAD

Hasta ahora se ha supuesto que el costo c de cada unidad es independiente del tamaño del pedido. Sin embargo, con frecuencia sucede que el proveedor está de acuerdo en cobrar menos por unidad cuando los pedidos son mayores. Esos descuentos por cantidad son comunes en muchos bienes de consumo.

Aunque existen muchas tarifas (planes, programas, etc.) de descuento de distintos tipos, hay dos que parecen las más usadas: sobre todas las unidades e incrementales. En cada caso se supone que hay uno o más puntos límite que definen los cambios en el costo unitario. Sin embargo, hay dos posibilidades: el descuento se aplica a todas las unidades en un pedido (totales), o sólo se aplica a las unidades adicionales respecto al punto límite (incrementales). El caso más común es el de descuento sobre todas las unidades.

Ejemplo *4.4* Weighty Trash Bags tiene la siguiente tabla de precios para sus bolsas grandes de basura. Para pedidos de menos de 500 bolsas, .30 dólares cada una; para más de 500 pero menos de 1 000, .29 dólares cada una, y para más de 1 000, .28 dólares cada una. En este caso, los puntos límite son 500 y 1 000. El esquema de descuento es sobre todas las unidades, porque el descuento se aplica a todas las unidades en un pedido. La función costo de pedido $C(Q)$ se define como sigue:

$$C(Q) = \begin{cases} 0.30Q & \text{para} \quad 0 \le Q < 500, \\ 0.29Q & \text{para} \quad 500 \le Q < 1\,000, \\ 0.28Q & \text{para} \quad 1\,000 \le Q \end{cases}$$

Esta función se ilustra en la figura 4-9. En la figura 4-10 se consideran los mismos puntos límite, pero se supone un esquema de descuento incremental.

Se observa que el costo promedio por unidad con un esquema de descuento en todas las unidades será menor que el correspondiente con el esquema incremental.

El esquema para todas las unidades parece irracional en ciertos aspectos. En el ejemplo anterior, 499 bolsas costarían 149.70 dólares, mientras que 500 sólo costarían 145 dólares. ¿Por

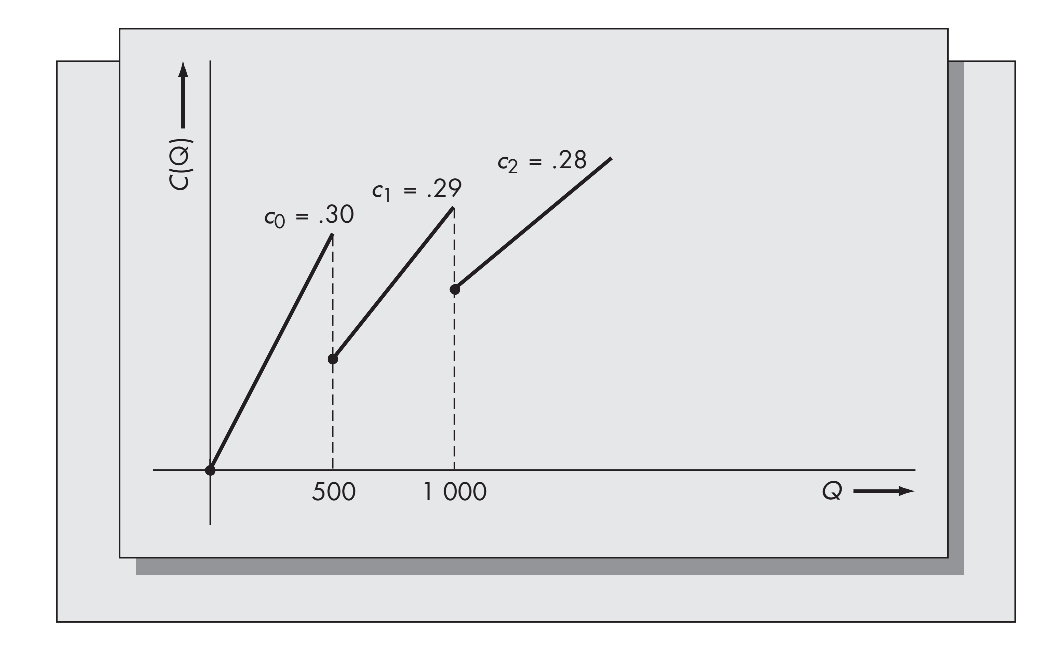


FIGURA 4-9
Función del costo de pedido con descuentos en todas las unidades.

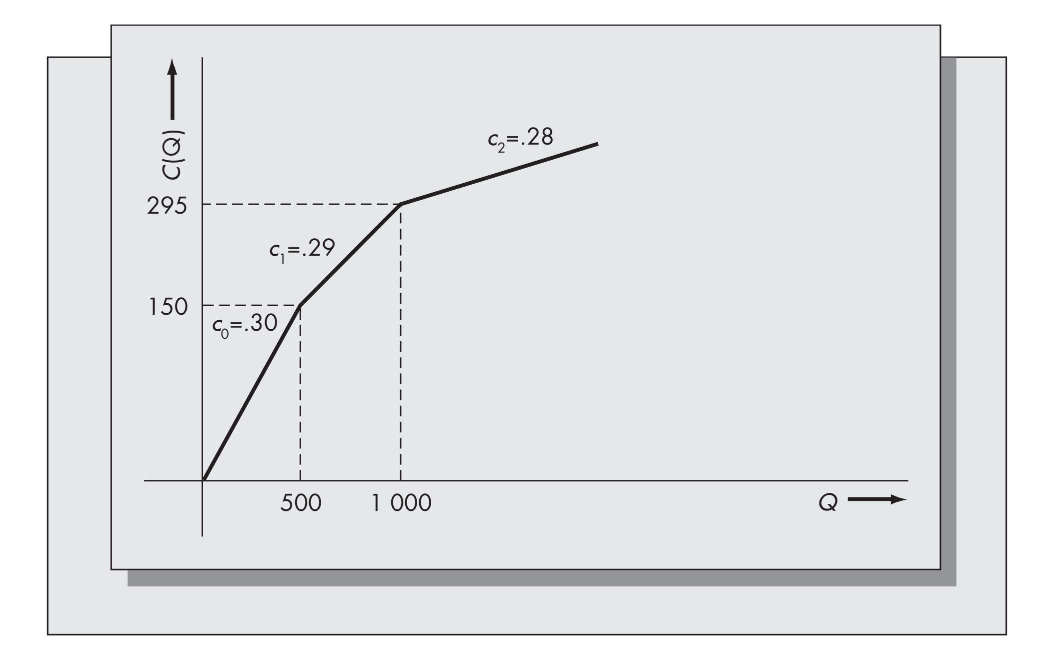


FIGURA 4-10
Función del costo de pedido con descuentos incrementales.

qué Weighty Trash Bags cobra en realidad menos en un pedido mayor? Una razón sería para dar un incentivo en la compra al mayoreo. Si usted quisiera comprar 400 bolsas, podría elegir llegar al punto límite para obtener el descuento. Además, es posible que Weighty Trash Bags haya almacenado sus productos en lotes de 100, de tal modo que sus ahorros en costos de manejo podrían más que compensar el menor costo total.

Política óptima para el esquema de descuento en todas las unidades

Ilustraremos la técnica de solución con el ejemplo 4.4 anterior. Supongamos que la empresa quiere saber qué pedido colocar con Weighty, y que su uso de bolsas de basura es relativamente constante (600 por año). El departamento de contabilidad estima que el costo fijo de hacer

un pedido es 8 dólares, y que los costos de inventario se basan en una tasa de interés anual de 20%. De acuerdo con el ejemplo 4.4, $c_0 = 0.30$, $c_1 = 0.29$ y $c_2 = 0.28$ son los respectivos costos unitarios.

El primer paso para llegar a la solución es calcular los valores de la cantidad óptima de pedido que corresponden a cada uno de los costos unitarios, que representaremos como $Q^{(0)}$, $Q^{(1)}$ y $Q^{(2)}$, respectivamente.

$$Q^{(0)} = \sqrt{\frac{2K\lambda}{Ic_0}} = \sqrt{\frac{(2)(8)(600)}{(0.2)(0.30)}} = 400,$$

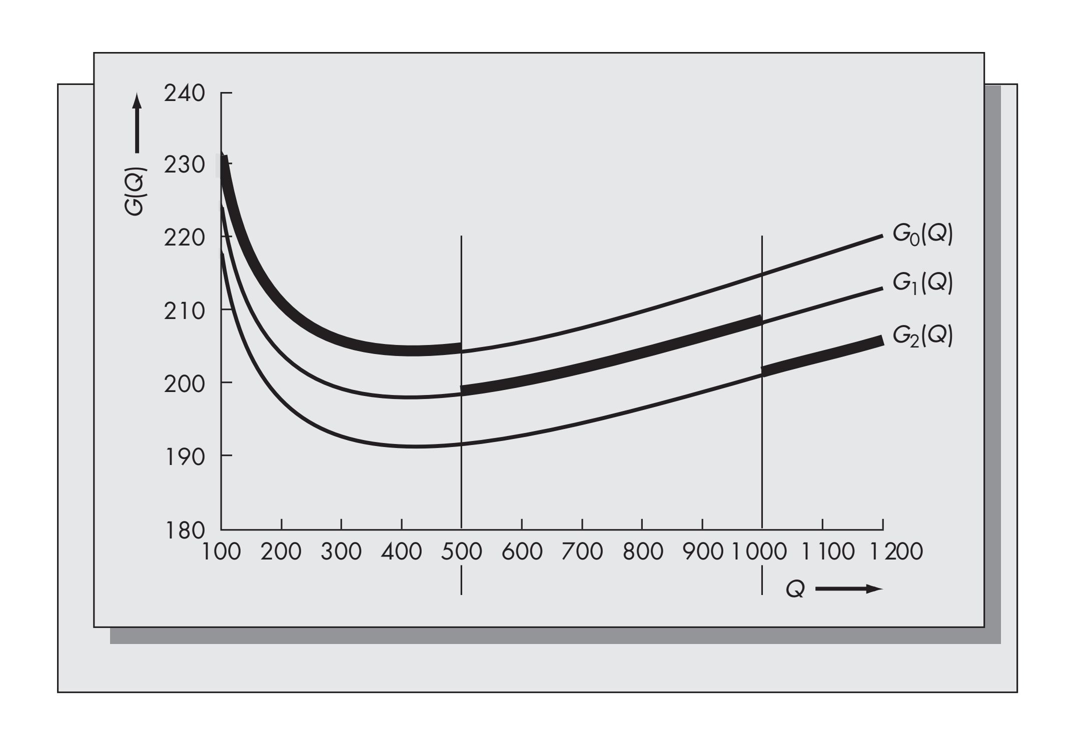
$$Q^{(1)} = \sqrt{\frac{2K\lambda}{Ic_1}} = \sqrt{\frac{(2)(8)(600)}{(0.2)(0.29)}} = 406,$$

$$Q^{(2)} = \sqrt{\frac{2K\lambda}{Ic_2}} = \sqrt{\frac{(2)(8)(600)}{(0.2)(0.28)}} = 414.$$

Se dice que el valor de CEP es realizable si cae dentro del intervalo que corresponde al costo unitario que se usa para calcularlo. Como $0 \leq 400 < 500$, Q^0 sí es realizable. Sin embargo, ni $Q^{(1)}$ ni $Q^{(2)}$ es realizable ($Q^{(l)}$ tendría que ser entre 500 y 1 000, y $Q^{(2)}$ debería ser 1 000 o más). Cada valor de cantidad económica de pedido corresponde al mínimo de una curva distinta de costo anual. En este ejemplo, si $Q^{(2)}$ fuera realizable, debería ser necesariamente la solución óptima, porque corresponde al punto mínimo de la curva más baja. Las tres curvas de costo anual promedio para este ejemplo se ilustran en la figura 4-11. Como cada curva sólo es válida para ciertos valores de Q, la función del costo anual promedio se representa con la curva discontinua gruesa. El objetivo del análisis es calcular el mínimo de esta curva discontinua.

Hay tres candidatos para la solución óptima: 400, 500 y 1 000. En general, la solución óptima será la CEP máxima realizable o uno de los puntos límite mayor que ella. Es el tamaño del lote con el mínimo costo anual promedio.

FIGURA 4-11
Función de costo anual con descuentos en todas las unidades.



Las funciones del costo anual promedio se expresan mediante

$$G_j(Q) = \lambda c_j + \lambda K/Q + Ic_j\, Q/2 \qquad \text{para } j = 0, 1, \text{ y } 2.$$

La curva discontinua de la figura 4-11, $G(Q)$, se define como sigue:

$$G(Q) = \begin{cases} G_0(Q) & \text{para} \quad 0 \le Q < 500, \\ G_1(Q) & \text{para} \quad 500 \le Q < 1\,000, \\ G_2(Q) & \text{para} \quad 1\,000 \le Q \end{cases}$$

Al sustituir $Q = 400$, 500 y 1 000, y usar los valores correspondientes de c_j, se obtiene

$$\begin{aligned} G(400) &= G_0(400) \\ &= (600)(0.30) + (600)(8)/400 + (0.2)(0.30)(400)/2 = \$204.00 \\ G(500) &= G_1(500) \\ &= (600)(0.29) + (600)(8)/500 + (0.2)(0.29)(500)/2 = \$198.10 \\ G(1\,000) &= G_2(1\,000) \\ &= (600)(0.28) + (600)(8)/1\,000 + (0.2)(0.28)(1\,000)/2 = \$200.80. \end{aligned}$$

Por lo anterior, se llega a la conclusión de que la solución óptima es hacer pedidos de 500 unidades con Weighty, a un costo anual promedio de 198.10 dólares.

Resumen de la técnica de solución para descuentos en todas las unidades

1. Determinar el valor máximo realizable de cantidad económica de pedido. La forma más eficiente de hacerlo es calcular la CEP primero para el precio mínimo y continuar con el siguiente más alto. Detenerse cuando el primer valor de CEP sea realizable (esto es, dentro del intervalo correcto).
2. Comparar el valor del costo promedio anual a la máxima CEP realizable y en todos los puntos límite de precios que sean mayores que la máxima CEP realizable. La Q óptima es el punto en el que se tiene el mínimo costo anual promedio.

Descuentos incrementales por cantidad

Regresemos al ejemplo 4.4, pero supongamos que los descuentos por cantidad son incrementales. Esto es, que las bolsas de basura cuestan 0.30 dólares cada una en cantidades de 500 o menos; para cantidades entre 500 y 1 000, las primeras 500 cuestan 0.30 dólares cada una y las restantes cuestan 0.29 dólares cada una, y para cantidades de 1 000 y más, las primeras 500 cuestan 0.30 dólares cada una, las siguientes 500 cuestan 0.29 dólares cada una y las restantes cuestan 28 centavos cada una. Se necesita deducir una ecuación matemática de la función $C(Q)$, que se ilustra en la figura 4-10. En ella se aprecia que la primera discontinuidad en el precio corresponde a $C(Q) = (500)(0.30) = 150$ dólares y la segunda corresponde a $C(Q) = 150 + (0.29 \text{ dólares})(500) = 295$ dólares. Por consiguiente,

$$C(Q) = \begin{cases} 0.30Q & \text{para} \quad 0 \le Q < 500, \\ 150 + 0.29(Q - 500) = 5 + 0.29Q & \text{para} \quad 500 \le Q < 1\,000, \\ 295 + 0.28(Q - 1\,000) = 15 + 0.28Q & \text{para} \quad 1\,000 \le Q \end{cases}$$

de modo que

$$C(Q)/Q = \begin{cases} 0.30 & \text{para} \quad 0 \le Q < 500, \\ 0.29 + 5/Q & \text{para} \quad 500 \le Q < 1\,000, \\ 0.28 + 15/Q & \text{para} \quad 1\,000 \le Q. \end{cases}$$

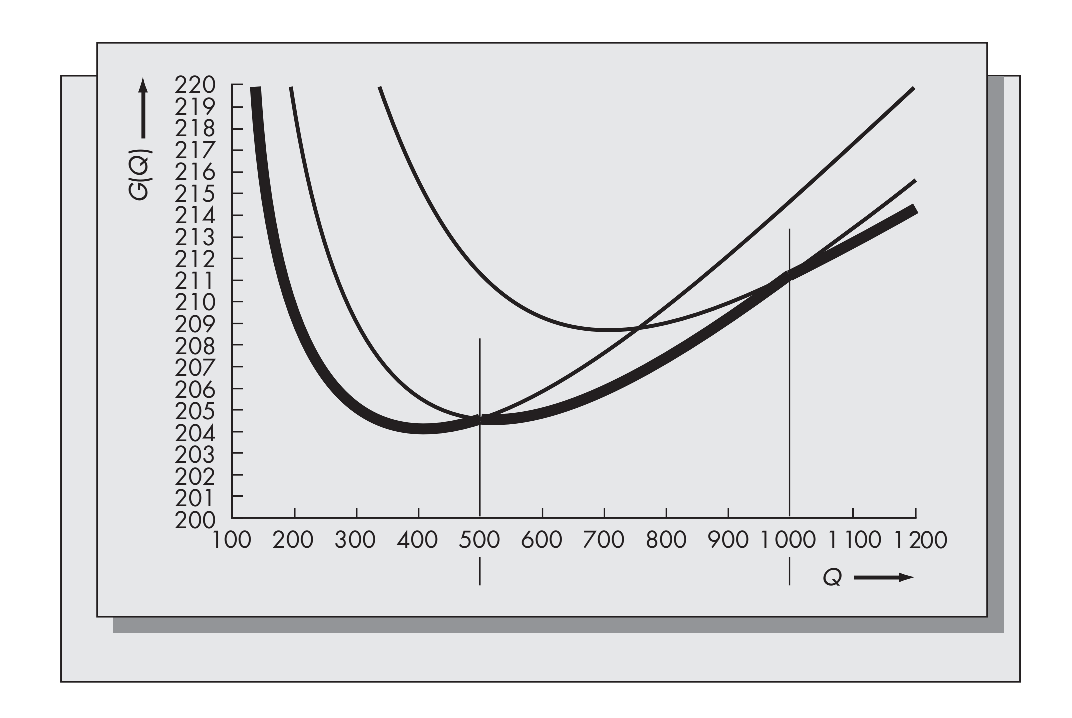


FIGURA 4-12
Función costo promedio anual para plan de descuento incremental.

La función del costo anual promedio, $G(Q)$, es

$$G(Q) = \lambda C(Q)/Q + K\lambda/Q + I[C(Q)/Q]Q/2.$$

En este ejemplo, $G(Q)$ tendrá tres representaciones algebraicas diferentes [$G_0(Q)$, $G_1(Q)$ y $G_2Q)$], dependiendo de en qué intervalo cae Q. Como $C(Q)$ es continua, $G(Q)$ también será continua. La función $G(Q)$ aparece en la figura 4-12.

La solución óptima se presenta en el mínimo de una de las curvas de costo anual promedio. Se obtiene sustituyendo las tres ecuaciones de $C(Q)/Q$ en la ecuación definitoria para $G(Q)$, calculando los tres mínimos de las curvas, determinando cuáles de esos mínimos caen en el intervalo correcto y, por último, comparando los costos anuales promedio en los valores realizables. Se ve que

$$G_0(Q) = (600)(0.30) + (8)(600)/Q + (0.20)(0.30)Q/2$$

que se minimiza en

$$Q^{(0)} = \sqrt{\frac{2K\lambda}{Ic_0}} = \sqrt{\frac{(2)(8)(600)}{(0.20)(0.30)}} = 400;$$

$$\begin{aligned} G_1(Q) &= (600)(0.29 + 5/Q) + (8)(600)/Q + (0.20)(0.29 + 5/Q)(Q/2) \\ &= (0.29)(600) + (13)(600)/Q + (0.20)(0.29)Q/2 + (0.20)(5)/2 \end{aligned}$$

que se minimiza en

$$Q^{(1)} = \sqrt{\frac{(2)(13)(600)}{(0.20)(0.29)}} = 519;$$

y finalmente

$$\begin{aligned} G_2(Q) &= (600)(0.28 + 15/Q) + (8)(600)/Q + (0.20)(0.28 + 15/Q)Q/2 \\ &= (0.28)(600) + (23)(600)/Q + (0.20)(0.28)Q/2 + (0.20)(15)/2 \end{aligned}$$

que se minimiza en

$$Q^{(2)} = \sqrt{\frac{(2)(23)(600)}{(0.20)(0.28)}} = 702.$$

Tanto $Q^{(0)}$ como $Q^{(1)}$ son realizables. $Q^{(2)}$ no es realizable, porque $Q^{(2)} < 1\,000$. La solución óptima se obtiene comparando $G_0(Q^{(0)})$ y $G_1(Q^{(1)})$. Si se sustituye en las ecuaciones anteriores para $G_0(Q)$ y $G_1(Q)$, se obtendrá

$$G_0(Q^{(0)}) = \$204.00,$$
$$G_1(Q^{(1)}) = \$204.58.$$

Por consiguiente, la solución óptima es colocar un pedido con Weighty por 400 unidades al precio máximo de 0.30 dólares por unidad. El costo de usar un pedido normal de 519 unidades sólo es un poco mayor. Observe que, comparado con el caso de todas las unidades, se obtiene un tamaño menor de lote a un costo anual promedio mayor.

Resumen de la técnica de solución para descuentos incrementales

1. Deducir una ecuación algebraica para $C(Q)$ correspondiente a cada intervalo de precio. Usarla para determinar una ecuación algebraica para $C(Q)/Q$.
2. Sustituir las ecuaciones deducidas para $C(Q)/Q$ en la ecuación definitoria para $G(Q)$. Calcular por separado el valor mínimo de Q que corresponde a cada intervalo de precio.
3. Determinar cuáles mínimos de los calculados en el paso 2 son realizables (esto es, caen en el intervalo correcto). Comparar los valores de los costos anuales promedio en los valores realizables de cantidad económica de pedido y elegir el mínimo.

Otros esquemas de descuento

Aunque los esquemas de descuento para todas las unidades e incremental son los más comunes, existen muchos otros. Un ejemplo es el esquema de descuento por furgón, representado en la figura 4-13.

La lógica en que se basa el descuento por furgón es la siguiente. Un furgón consiste en M unidades. El proveedor cobra c constante por unidad hasta haber pagado el costo de un furgón

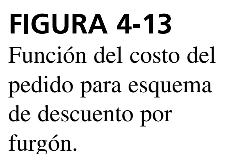
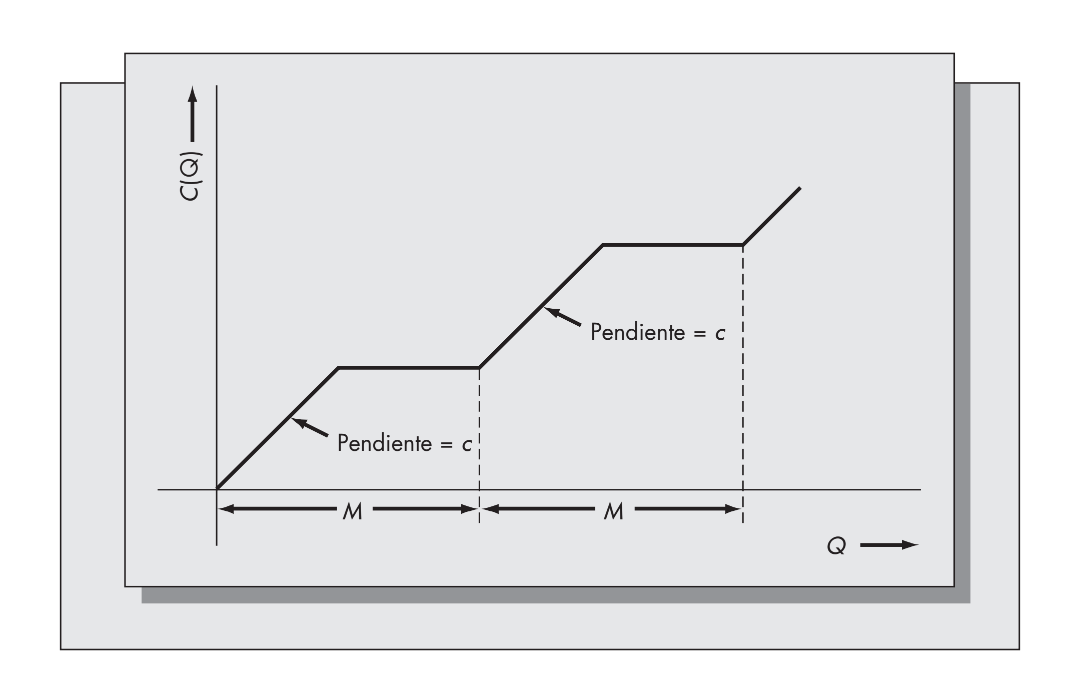


FIGURA 4-13
Función del costo del pedido para esquema de descuento por furgón.



completo, y en ese punto ya no se cobran las unidades adicionales de ese furgón. Una vez que se llena el primer furgón, se paga c por unidad hasta llenar el segundo furgón, y así sucesivamente.

Determinar las políticas óptimas para el plan de descuento por furgón, u otros esquemas, puede ser extremadamente difícil. Cada esquema de descuento tiene una solución exclusiva. Algunos procedimientos pueden tener una complejidad extrema.

Problemas para la sección 4.7

21. La tienda local de víveres almacena rollos de toallas para baño en paquetes individuales y en paquetes de 12 toallas, más económicos. Usted trata de decidir cuál comprar. El paquete individual cuesta 0.45 dólares y el de una docena cuesta 5.00 dólares. Usted consume las toallas de papel a una tasa relativamente constante de un rollo cada tres meses. Su costo de oportunidad se calcula suponiendo una tasa de interés de 25% y un costo fijo de 1 dólar por el tiempo adicional que le toma ir a la tienda a comprar el papel. (Suponemos que usted va de compras a la tienda con la suficiente frecuencia para no requerir un viaje especial cuando se le agotan las toallas.)

    a) ¿Cuántos rollos sencillos compraría usted para minimizar los costos anuales de inventario y de preparación al comprar toallas para baño?

    b) Determine si es más económico comprar las toallas en paquetes de 12.

    c) ¿Hay motivos adicionales a los que se describieron en el problema que pudieran motivar el no adquirir los paquetes de 12 rollos?

22. Un agente de compras de determinada oblea de silicio que se emplea en la producción de semiconductores debe optar por uno de tres proveedores. El proveedor A le vende las obleas a 2.50 dólares cada una, independientemente de la cantidad pedida. El proveedor B se las vende a 2.40 dólares cada una, pero no recibe pedidos menores de 3 000 obleas, y el proveedor C se las vende a 2.30 dólares por pieza, pero no acepta pedidos menores de 4 000 obleas. Suponga que el costo de preparación de pedido es 100 dólares, y que el consumo anual son 20 000 obleas. Suponga una tasa de interés anual de 20% para sus cálculos del costo de mantener el inventario.

    a) ¿Qué proveedor debe usar, y cuál es el tamaño del pedido?

    b) ¿Cuál es el valor óptimo de los costos de inventario y preparación de las obleas cuando se acude al proveedor óptimo?

    c) Si el tiempo de retraso para reabastecimiento de obleas es tres meses, calcule el punto de reorden basado en el nivel disponible de inventario de obleas.

23. Suponga que han pasado dos años, y que el agente de compras del problema 22 debe recalcular la cantidad óptima de obleas por comprar y a qué proveedor deben comprarse. El proveedor B ha decidido aceptar las ofertas de cualquier tamaño, pero vende sus obleas a 2.55 dólares cada una en pedidos de hasta 3 000, y a 2.25 dólares por pieza en cantidades superiores a 3 000 obleas. El proveedor A todavía conserva su esquema de precios, y el proveedor C salió del mercado. Ahora, ¿cuál proveedor se debe elegir?

24. En el cálculo de una política óptima para un esquema de descuento en todas las unidades, usted calcula primero los valores de las CEP para cada uno de los costos con tres intervalos, y obtiene: $Q^{(0)} = 800$, $Q^{(1)} = 875$ y $Q^{(2)} = 925$. Ese plan tiene puntos límite en 750 y en 900. Con base sólo en esta información, ¿podría determinar cuál es la cantidad óptima de pedido? Explique su respuesta.

25. Sistemas Parasol vende tarjetas para computadoras personales. En cantidades de hasta 25 tarjetas vende a 350.00 dólares cada pieza; en cantidades adicionales, cobra 315 dólares por cada tarjeta, desde 26 hasta 50, y a 280 dólares cada tarjeta adicional cuando se adquieran más de 50. Una gran empresa de comunicaciones considera que necesitará esas tarjetas durante los 10 años siguientes, y su consumo será al menos de 140 por año. Los costos de preparación de pedido son 30 dólares, y los de mantener el inventario se basan en la tasa anual de interés de 18%. ¿Cuál debe ser el tamaño del pedido?

*4.8 SISTEMAS DE MÚLTIPLES PRODUCTOS CON RESTRICCIONES DE RECURSOS

El modelo EOQ y sus extensiones aplican únicamente para artículos de inventario. Sin embargo, estos modelos se utilizan en compañías que almacenan diferentes artículos. A pesar de que se pueden registrar órdenes por cantidades separadas para cada uno de los diferentes artículos, existen limitaciones que causarían soluciones infactibles. En el ejemplo 4.1, la solución óptima sugiere una compra de 3 870 lápices cada 1.24 años. La tienda, en cambio, puede carecer de espacio suficiente para almacenar tal cantidad de lápices, o suficiente dinero para adquirirlos todos en una sola compra.

Ejercicio *4.5*

En un pequeño taller de fabricación se producen tres artículos. La administración de ese taller ha establecido el requisito de nunca tener más de 30 000 dólares invertidos en inventario de esos artículos en determinado momento. Emplea una tasa anual de interés de 25% para calcular el costo de mantener el inventario. Los parámetros relevantes de costo y demanda están registrados en la siguiente tabla. ¿Qué tamaños de lote debe producir el taller para no exceder el presupuesto?

	Artículo		
	1	**2**	**3**
Tasa de demanda λ_j	1 850	1 150	800
Costo variable c_j	50	350	85
Costo de preparación K_j	100	150	50

Solución

Si el presupuesto no se excede al usar los valores de cantidad óptima de pedido para los tres artículos, las CEP serán óptimas. Por consiguiente, el primer paso es calcular los valores de CEP para todos los artículos y poder determinar si se viola la restricción.

$$\text{EOQ}_1 = \sqrt{\frac{(2)(100)(1\,850)}{(0.25)(50)}} = 172,$$

$$\text{EOQ}_2 = \sqrt{\frac{(2)(150)(1\,150)}{(0.25)(3\ \ 0)}} = 63,$$

$$\text{EOQ}_3 = \sqrt{\frac{(2)(50)(800)}{(0.25)(85)}} = 61.$$

Si se usa el valor CEP de cada artículo, la inversión máxima en inventario sería

(172)(50) + (63)(350) + (61)(85) = $35 835.[3]

Como la solución de la cantidad económica de pedido viola la restricción, necesitamos reducir esos tamaños de lote. Pero, ¿cómo?

[3] Estamos suponiendo para fines de esta sección que los tres pedidos se han hecho simultáneamente. Al escalonarse los ciclos de pedido, es posible cumplir con la restricción con tamaños de lote más grandes. Ese caso no lo consideramos aquí.

En este caso la solución óptima es muy fácil de calcular. Tan sólo se multiplica cada valor de CEP por la relación (30 000 dólares)/(35 835 dólares) = 0.8372. Para tener la garantía de no exceder el presupuesto de 30 000 dólares, se redondea cada valor hacia abajo (se pueden hacer ajustes después). Si Q_1^*, Q_2^* y Q_3^* son los valores óptimos, entonces

$$Q_1^* = (172)(0.8372) \approx 144,$$
$$Q_2^* = (63)(0.8372) \approx 52,$$
$$Q_3^* = (61)(0.8372) \approx 51,$$

en donde el signo $\approx$ debe interpretarse como redondeo al siguiente entero inferior en este caso.

El presupuesto total requerido para esos tamaños de lote es 29 735 dólares. Los 265 dólares restantes se pueden usar ahora para aumentar un poco los tamaños de lote de los productos 1 y 3. (Por ejemplo, $Q_1^* = 147$, $Q_3^* = 52$ dan como resultado un presupuesto de 29 970 dólares.)

En general, los problemas de restricción de presupuesto o de espacio no se resuelven con tanta facilidad. Suponga que hay n artículos cuyos costos unitarios son $c_1, c_2, \ldots, c_n$, respectivamente, y que el presupuesto total disponible para ellos es C. Entonces la restricción del presupuesto se puede escribir de la siguiente manera:

$$c_1Q_1 + c_2Q_2 + \cdots + c_nQ_n \leq C.$$

Sea

$$\text{CEP}_i = \sqrt{\frac{2K_i\lambda_i}{h_i}} \qquad \text{para } i = 1, \ldots, n,$$

siendo K_i, h_i y λ_i los parámetros respectivos de costo y demanda.

Hay dos posibilidades: la restricción se viola o no (*es activa o no*). Si no es activa, entonces

$$\sum_{i=1}^{n} c_i\text{CEP}_i \leq C,$$

y la solución óptima es $Q_i = \text{CEP}_i$. Si la restricción es activa, entonces

$$\sum_{i=1}^{n} c_i\text{CEP}_i > C,$$

y la solución de CEP no es factible. Si se incluye el siguiente supuesto, la solución en este caso es relativamente fácil de calcular:

$$\textit{Suposición: } c_1/h_1 = c_2/h_2 = \cdots = c_n/h_n.$$

Si ese supuesto es válido y la restricción es activa, en el apéndice 4-A se demuestra que la solución óptima es

$$Q_i^* = m\text{CEP}_i,$$

en la que el multiplicador m se calcula con

$$m = C\Big/\left[\sum_{i=1}^{n} (c_i\text{EOQ}_i)\right].$$

Como $c_i/h_i = c_i/(I_ic_i) = 1/I_i$, la condición de que las relaciones sean iguales equivale a usar la misma tasa de interés al calcular el costo de mantener el inventario para cada artículo, que es razonable en la mayoría de las circunstancias.

Suponga que la restricción es de espacio disponible. Sea w_i el espacio ocupado por una unidad del producto i, para $i = 1, 2, \ldots, n$ (podría ser espacio de piso expresado en pies cuadra-

dos, o volumen en pies cúbicos), y sea W el espacio total disponible. Entonces la restricción del espacio tiene la forma

$$w_1Q_1 + w_2Q_2 + \cdots + w_nQ_n \leq W.$$

Esto, matemáticamente hablando, tiene la misma forma que la restricción de presupuesto, así que se le aplica el mismo análisis. Sin embargo, la condición para tener una solución sencilla en este caso es que las relaciones w_i/h_i sean iguales. Esto es, que el espacio ocupado por un artículo debe ser proporcional a su costo de inventario. Cuando la tasa de interés es fija, esto equivale al requisito de que el espacio ocupado sea proporcional al valor del artículo. Dicho requisito probablemente sea demasiado restrictivo en la mayoría de los casos. Por ejemplo, las plumas fuente ocupan mucho menos espacio que los paquetes de papel oficio, pero son más costosas.

Examinemos ahora el problema en el que la restricción es activa, pero el supuesto de proporcionalidad no se cumple. Este problema es mucho más complicado que el que resolvimos antes. Requiere la formulación de una función *lagrangiana.* Los detalles de esa formulación para este problema se pueden ver en el apéndice 4-A. Como se demostró, los tamaños óptimos tienen ahora la forma

$$Q_i^* = \sqrt{\frac{2K_i\lambda_i}{h_i + 2\theta w_i}},$$

donde θ es una constante que se elige para que

$$\sum_{i=1}^{n} w_iQ_i^* = W.$$

La constante θ, conocida como multiplicador de Lagrange, reduce los tamaños de lote aumentando el costo efectivo de inventario. Es posible determinar el valor de θ mediante tanteos o con una técnica de búsqueda como bisección de intervalo. Observe que $\theta > 0$, de modo que la búsqueda puede limitarse sólo a números positivos. Se puede interpretar el valor de θ como la disminución en el costo promedio anual que resultaría al agregar una unidad más de recurso. En este caso, representa el beneficio marginal de un pie cuadrado más de espacio.

Ejemplo *4.6*

Regresemos al taller de fabricación del ejemplo 4.5. Además de la restricción del presupuesto, supongamos que sólo hay 2 000 pies cuadrados de piso disponibles, y que los tres productos ocupan, respectivamente, 9, 12 y 18 pies cuadrados por unidad.

Primero hay que comprobar que la solución CEP sea factible. Al igualar $w_i = 9$, $w_2 = 12$ y $w_3 = 18$, encontramos que

$$\sum \text{CEP}_i w_i = (172)(9) + (63)(12) + (61)(18) = 3\,402,$$

lo cual obviamente no es factible. A continuación se comprueba si la solución restringida por presupuesto produce una solución factible al problema restringido por espacio. Para la solución restringida por presupuesto se requiere $(147)(9) + (52)(12) + (52)(18) = 2\,883$ pies cuadrados de espacio, lo cual tampoco es factible.

El siguiente paso es calcular las relaciones w_i/h_i para $1 \leq i \leq 3$. Se ve que esas relaciones son 0.72, 0.14 y 0.85, respectivamente. Como sus valores son distintos, la solución simple obtenida con un escalamiento proporcional de los valores de CEP no será óptima. Por consiguiente, debe calcularse el valor de θ, el multiplicador de Lagrange.

Podemos determinar las cotas superior e inferior del valor óptimo de θ suponiendo que las relaciones son iguales. Si fueran iguales, el multiplicador, m, sería

$$m = W\Big/\sum(\text{CEP}_i w_i) = 2\,000/3\,402 = 0.5879,$$

que daría como resultado los tres tamaños de lote de 101, 37 y 36, respectivamente. Los tres valores de θ que dan como resultado esos tamaños de lote son, respectivamente, $\theta = 1.32$, $\theta = 6.86$ y $\theta = 2.33$. (Esos valores se obtuvieron sustituyendo los tamaños de lote 101, 37 y 36 en la ecuación anterior de Q_1^* y despejando θ.)

El valor real de θ estará entre 1.32 y 6.86. Si comenzamos con $\theta = 3.5$, obtendremos $Q_1^* = 70$, $Q_2^* = 45$ y $Q_3^* = 23$, y $\Sigma w_i Q_i^* = (70)(9) + (45)(12) + (23)(18) = 1\,584$, que significa que $\theta < 3.5$. Después de bastantes pruebas se llega finalmente al valor óptimo de $\theta = 1.75$ y $Q_1^* = 92$, $Q_2^* = 51$ y $Q_3^* = 31$, que dan $\Sigma w_i Q_i^* = 1\,998$. Observe que estos tamaños de lote tienen proporciones muy distintas a las obtenidas suponiendo un multiplicador constante. La búsqueda del valor óptimo del multiplicador de Lagrange, aunque tediosa cuando se realiza a mano, puede efectuarse eficientemente con computadora. Las hojas de cálculo son métodos adecuados para hacer esas operaciones.

Tan sólo hemos tocado algunas de las complicaciones que se pueden presentar al aplicar modelos tipo cantidad económica de pedido a un sistema real con múltiples productos. Con frecuencia los problemas reales son demasiado complicados para poder expresarlos con precisión en forma de un modelo matemático soluble. Por esto, los modelos sencillos como los presentados en este capítulo y en los siguientes se utilizan en la práctica. Sin embargo, cualquier solución que recomiende un modelo matemático debe examinarse en el contexto del sistema en el que se usará.

Problemas para la sección 4.8

26. Un puesto de verduras tiene exactamente 1 000 pies cuadrados de espacio para presentar tres productos: tomates, lechugas y calabazas. Los datos correspondientes a cada uno se ven en la tabla siguiente:

	Artículo		
	Tomates	**Lechugas**	**Calabazas**
Demanda anual (en libras)	850	1 280	630
Costo por libra	$0.29	$0.45	$0.25

El costo de preparación para reabastecer las verduras es 100 dólares en cada caso, y el espacio ocupado por cada verdura es proporcional a sus costos. Los tomates requieren .5 pies cuadrados por libra. La tasa de interés anual que se usa para calcular los costos de inventario es 25%. ¿Cuáles son las cantidades óptimas de las tres verduras que se deben comprar?

27. Suponga que las verduras del problema 26 se compran en momentos distintos. ¿En qué forma puede afectar eso la política de compras del dueño del puesto?

28. Suponga que en el problema 26 el espacio ocupado por cada verdura no es proporcional a su costo. Suponga que una libra de lechuga requiere 4 pies cuadrados de piso y que 1 libra de calabazas requiere 1 pie cuadrado de piso. Calcule las cotas superior e inferior de las cantidades óptimas de pedido en este caso. Ensaye con distintos valores del multiplicador de Lagrange para calcular los valores óptimos de las cantidades de pedido. (Una hoja de cálculo es ideal para este tipo de cálculos. Si resuelve el problema con una hoja de cálculo, ponga el multiplicador de Lagrange en una celda para poder cambiar con facilidad su valor.)

4.9 MODELOS DE CANTIDAD ÓPTIMA DE PEDIDO PARA PLANEACIÓN DE LA PRODUCCIÓN

Los modelos sencillos de dimensionamiento de lotes se han aplicado muy bien a una diversidad de problemas de manufactura. En esta sección se describe una extensión del modelo CEP con una tasa de producción finita, señalada en la sección 4.6, al problema de fabricar n productos en una sola máquina.

De acuerdo con la notación usada en este capítulo, sean

λ_j = Tasa de demanda para el producto j.

P_j = Tasa de producción para el producto j.

h_j = Costo de inventario por unidad de producto j y por unidad de tiempo.

K_j = Costo de preparar la instalación productiva para fabricar el producto j.

El objetivo es determinar el procedimiento óptimo para procesar n productos en la máquina de tal manera que se minimice el costo de inventario y de preparación, y garantizar que no haya agotamientos de inventario durante el ciclo de producción.

Se requiere suponer que

$$\sum_{j=1}^{n} \lambda_j/P_j \leq 1.$$

Este supuesto es necesario para asegurar que la instalación tenga la capacidad suficiente para satisfacer la demanda de todos los productos. Observe que es más estricto que el supuesto hecho en la sección 4.6, que $\lambda_j < P_j$ para cada j. Para visualizar por qué es necesario este supuesto, imaginemos el caso de dos productos con tasas idénticas de demanda y producción. En cada ciclo se fabricaría primero el producto 1 y después el producto 2. Es claro que $P_1 \geq 2\lambda_1$ y $P_2 \geq 2\lambda_2$, de modo que se puede producir bastante para cumplir con la demanda total de ambos productos cada ciclo. Esto se traduce en que $\lambda_1/P_1 \leq 0.5$ y $\lambda_2/P_2 \leq 0.5$, obteniendo un valor de la suma menor o igual a 1. Un razonamiento semejante es válido para más de dos productos con tasas de demanda y producción no idénticas.

También supondremos que la política adoptada es *política de ciclo de rotación.* Esto significa que en cada ciclo hay exactamente una preparación para cada producto, y los productos se fabrican en la misma secuencia en cada ciclo de producción. La importancia de este supuesto se describe al final de la sección.

En un inicio cabría pensar que la solución óptima es producir secuencialmente tamaños de lote para cada producto, optimizados considerando cada producto por separado. De acuerdo con la sección 4.6, se obtendría así un tamaño de lote del producto j igual a

$$Q_j = \sqrt{\frac{2K_j\lambda_j}{h'_j}},$$

en donde $h'_j = h_j(1 - \lambda_j/P_j)$. El problema con este método es que, como sólo se cuenta con un equipo de producción, es probable que algunos de los tamaños de lote Q_j no sean suficientemente grandes para cumplir con la demanda entre las corridas de fabricación del producto j y por consiguiente los inventarios se agoten.

Sea T el tiempo del ciclo. Durante el tiempo T se supone que se fabrica exactamente un lote de cada producto. Para que el lote del producto j sea suficientemente grande para cumplir con la demanda que se presente durante el tiempo T, el tamaño de lote debe ser

$$Q_j = \lambda_j T.$$

De acuerdo con la sección 4.6, el costo anual promedio asociado con el producto j se puede expresar así:

$$G(Q_j) = K_j\lambda_j/Q_j + h'_jQ_j/2.$$

El promedio anual de costo para todos los productos es la suma

$$\sum_{j=1}^{n} G(Q_j) = \sum_{j=1}^{n} K_j\lambda_j/Q_j + h'_jQ_j/2.$$

Al sustituir $T = Q_j/\lambda_j$ se obtiene el costo promedio anual asociado con los n productos en términos del tiempo del ciclo T, que es

$$G(T) = \sum_{j=1}^{n} [K_j/T + h'_j\lambda_j T/2].$$

El objetivo es determinar a T que minimiza a $G(T)$. La condición necesaria para un T óptimo es

$$\frac{dG(T)}{dT} = 0.$$

Igualando a cero la primera derivada con respecto a T se obtiene

$$\sum_{j=1}^{n} [-K_j/T^2 + h'_j\lambda_j/2] = 0.$$

Al despejar T se obtiene el tiempo óptimo de ciclo T^*

$$T^* = \sqrt{\frac{2\sum_{j=1}^{n} K_j}{\sum_{j=1}^{n} h'_j\lambda_j}}.$$

Si los tiempos de preparación son relevantes, debe comprobarse que haya suficiente tiempo cada ciclo para tener en cuenta los tiempos de preparación y producción de los n productos. Sea s_j el tiempo de preparación para el producto j. Para asegurar que el tiempo total requerido para las preparaciones y la producción en cada ciclo no rebase T se obtiene la restricción

$$\sum_{j=1}^{n} (s_j + Q_j/P_j) \le T.$$

Aprovechando que $Q_j = \lambda_j T$, esto se traduce en

$$\sum_{j=1}^{n} (s_j + \lambda_j T/P_j) \le T,$$

que da como resultado, después de reordenar términos,

$$T \ge \frac{\sum_{j=1}^{n} s_j}{1 - \sum_{j=1}^{n} (\lambda_j/P_j)} = T_{\text{mín}}.$$

Como $T_{\text{mín}}$ no se puede rebasar sin comprometer la factibilidad, la solución óptima es elegir el tiempo de ciclo T igual al valor *mayor* de entre T^* y $T_{\text{mín}}$.

Ejemplo *4.7*

Bali produce varios estilos de calzado para dama y caballero en una instalación cercana a Bérgamo, Italia. La piel para las cubiertas y suelas se corta en una sola máquina. Esta planta de Bérgamo fabrica siete estilos y varios colores de cada uno. (Para nuestros fines, no se consideran productos distintos los diversos colores, porque no se requiere preparación al cambiar color.) A Bali le gustaría programar el corte de los zapatos recurriendo a una política de rotación que cumpla con toda la demanda y minimice los costos de preparación e inventario. Los costos y tiempos de preparación son proporcionales.

TABLA 4-1 Datos relevantes para el ejemplo 4.7

Estilo	Demanda anual (unidades/año)	Tasa de producción (unidades/año)	Tiempo de preparación (horas)	Costo variable (dólares/unidad)
Zapatilla para dama	4 520	35 800	3.2	40
Pantufla para dama	6 600	62 600	2.5	26
Bota para dama	2 340	41 000	4.4	52
Sandalia para dama	2 600	71 000	1.8	18
Mocasín para caballero	8 800	46 800	5.1	38
Pantufla para caballero	6 200	71 200	3.1	28
Zapato para caballero	5 200	56 000	4.4	31

La empresa estima que ascienden a un promedio de 110 dólares por hora, con base en el costo del tiempo de operador y el del tiempo ocioso de las máquinas durante las preparaciones. Los costos de inventario se basan en una tasa de interés anual de 22%.

Los datos relevantes para este problema se asientan en la tabla 4-1.

Solución

El primer paso es comprobar que el problema sea factible. Para ello se calcula $\Sigma\, \lambda_j/P_j$. El lector debe comprobar que esa suma es igual a 0.69355. Como es menor que 1, habrá una solución factible. A continuación se calcula el valor de T^*, pero para hacerlo se necesitan varios cálculos intermedios.

Primero se calculan los costos de preparación, que se supone son 110 dólares por los tiempos de preparación. A continuación se calculan los costos modificados de mantener el inventario (h_j'). Para ello se multiplica el costo de cada producto por la tasa anual de interés (0.22) y por el factor $1 - \lambda_j/P_j$. Los resultados son los siguientes:

Costos de preparación (K_j)	Costos modificados de mantener el inventario (h_j')
352	7.69
275	5.12
484	10.79
198	3.81
561	6.79
341	5.62
484	6.19

La suma de los costos de preparación es 2 695 dólares, y la de los productos de los costos modificados de inventario por las demandas anuales es 230 458.4 dólares. Al sustituir estas cifras en la fórmula de T^* se obtiene el tiempo óptimo del ciclo igual a 0.1529 años. Suponiendo un año con 250 días laborables, esto significa que el ciclo de rotación debe repetirse aproximadamente cada 38 días de trabajo. El tamaño óptimo del lote para cada tipo de zapato se calcula multiplicando el tiempo del ciclo por la tasa de la demanda para cada artículo. El lector debe comprobar que resultan los siguientes tamaños de lote:

Estilo	Tamaños óptimos de lote por cada corrida de producción
Zapatilla para dama	691
Pantufla para dama	1 009
Bota para dama	358
Sandalia para dama	398
Mocasín para caballero	1 346
Pantufla para caballero	948
Zapato para caballero	795

La planta cortaría las suelas y las cubiertas para estos tamaños de lote en orden (aunque la secuencia no necesariamente deba ser la de la tabla) y repetiría el ciclo de rotación cada 38 días (0.1529 año).

Sin embargo, esta solución sólo puede implementarse si T^* no es menor que $T_{\text{mín}}$. Para calcular $T_{\text{mín}}$ debemos expresar los tiempos de preparación en años. Suponiendo 8 horas de trabajo por día y 250 días laborables por año, se dividirían los tiempos de preparación en horas entre 2 000 (250 por 8). El lector debe comprobar que el valor resultante de $T_{\text{mín}}$ sea 0.04; con ello se ve que T^* es factible y, por consiguiente, óptimo.

El costo total anual promedio de mantener inventario y preparaciones con la política óptima se puede determinar calculando el valor de $G(T)$ cuando $T = T^*$. En este caso es 35 244.44 dólares. Es interesante notar que si el gerente de la planta opta por implementar esta política, la instalación quedará ociosa durante una parte considerable de cada ciclo de rotación. El tiempo total de subida en cada ciclo de rotación se calcula dividiendo los tamaños de lote por las tasas de producción para cada estilo y sumando los resultados. Así se llega a 0.106 años. Por consiguiente, la política óptima de rotación que minimiza los costos totales de inventario y de preparación da como resultado que la operación de corte permanece ociosa aproximadamente la tercera parte del tiempo.

El lector debe percatarse de que la solución relativamente sencilla a este problema fue consecuencia de dos supuestos. Uno fue que los tiempos de preparación no dependían de la secuencia. En el ejemplo anterior es posible que el tiempo requerido para cambiar de un estilo de calzado a otro dependa de los estilos. Por ejemplo, un cambio de un estilo para dama a uno para caballero probablemente sea más tardado que entre dos para dama. Un segundo supuesto fue que la planta tenía una política de ciclo de rotación. Esto significa que Bali hace en cada ciclo una sola corrida de producción para cada estilo. Cuando las tasas de demanda y los costos de preparación difieren mucho, podría ser mejor tener dos o más corridas de producción de un producto en un ciclo. Hasta donde se sabe, este problema, más general, no se ha resuelto. En Maxwell (1964) se puede ver una discusión de estos asuntos y otros relacionados. Magee y Boodman (1967) presentan algunas formas heurísticas de manejar el problema de manera más general.

Problemas para la sección 4.9

29. Un taller de manufactura metálica tiene una sola troqueladora. En la actualidad hay un convenio para producir tres partes que requiere la prensa y parece que éstas seguirán fabricándose en el futuro. Se puede suponer que la prensa es el recurso crítico para esas partes, y que no se necesita considerar la interacción entre la prensa y las demás máquinas en el taller. La información relevante en este caso es la siguiente:

Número de parte	Cantidad anual contratada (demanda)	Costo de preparación (dólares)	Costo en dólares (por unidad)	Tasa de producción (anual)
1	2 500	80	16	45 000
2	5 500	120	18	40 000
3	1 450	60	22	26 000

Los costos de mantener el inventario se basan en una tasa de interés anual de 18%, y se fabricarán los productos en secuencia en un ciclo de rotación. Puede considerarse que los tiempos de preparación son despreciables.

a) ¿Cuál es el tiempo óptimo entre preparaciones para la parte 1?

b) ¿Durante qué porcentaje del tiempo está ociosa la troqueladora, suponiendo una política de ciclo óptimo de rotación?

c) ¿Cuáles son los tamaños óptimos de lote para cada parte que procesa la prensa en una solución óptima?

d) ¿Cuál es el costo total anual de inventario y preparación para estos artículos en la troqueladora, suponiendo un ciclo óptimo de rotación?

30. Tomlinson Furniture tiene un solo torno para trabajar la madera, que se emplea para elaborar partes como barrotes, patas, etc. En el torno se tallan cuatro formas y se producen lotes que pasan a inventario. Para simplificar la programación se produce un lote de cada tipo en un ciclo, y éste puede comprender tiempo ocioso. Los cuatro productos y su información relevante aparecen en la siguiente tabla.

Pieza	Necesidades mensuales	Tiempo de preparación (horas)	Costo unitario (dólares)	Tasa de producción (unidades/día)
J-55R	125	1.2	20	140
H-223	140	0.8	35	220
K-18R	45	2.2	12	100
Z-344	240	3.1	45	165

El tiempo de operador para las preparaciones se valúa en 85 dólares por hora, y los costos de mantener el inventario se basan en una tasa anual de 20% de interés. Suponga 20 días de trabajo por mes y 12 meses por año en sus cálculos.

a) Calcule la longitud óptima del ciclo de rotación.

b) ¿Cuáles son los tamaños óptimos de lote para cada producto?

c) ¿Cuáles son los porcentajes de tiempo de subida y de bajada para el torno, suponiendo que no se use para otros trabajos?

d) Trace una gráfica que muestre el cambio en el nivel de inventario durante un ciclo normal para cada producto.

e) Comente por qué la solución que obtuvo podría no ser factible para la empresa, o por qué podría no ser deseable aunque sí fuera factible.

4.10 PROGRAMAS INFORMÁTICOS PARA CONTROL DE INVENTARIOS

Antes de la proliferación de la computadora personal en la década de los ochenta sólo había algunos programas disponibles para aplicaciones de control de inventarios. En muchos casos los fabricantes de los sistemas de cómputo, como IBM, Control Data, DEC y otros vendían estos productos. El sistema prototipo de control de inventario para computadoras centrales era IMPACT de IBM (supuestamente con la colaboración de R. G. Brown). IMPACT usaba suavizamiento exponencial para rastrear la demanda (como se describe en el capítulo 2) y establecía puntos de reorden basados en la distribución de los errores de pronóstico (descritos en los capítulos 2 y 5). Desde entonces IBM ha desarrollado y comercializado algunas variantes de IMPACT para mercados específicos. Un ejemplo es INFOREM, diseñado para ventas al menudeo.

Hasta donde sabe este autor, los sistemas de control de inventario que vende IBM en la actualidad están diseñados sólo para computadoras centrales. Como el costo de las computadoras personales ha caído tan drásticamente en los últimos años y su poder ha mejorado constantemente, casi todas las empresas pequeñas los usan. Esto significa que cualquier persona tiene acceso al control computarizado de inventarios.

La mayoría de los programas de control de inventarios para computadora personal son módulos de un sistema integrado de contabilidad. Algunos de esos paquetes de contabilidad son costosos, y su precio supera en ocasiones al de las computadoras. En forma característica, los programas sencillos sólo permiten funciones básicas de mantenimiento de registros. Los modelos que se describieron en el capítulo 2 para pronósticos y control de inventario que se describen aquí y en el capítulo 5, sólo se incorporan en los mejores paquetes. Algunos productos también respaldan métodos justo a tiempo de administración de inventarios.

Una descripción detallada de los programas de control de inventario para computadoras personales fue publicado por Cohn (1994). Los productos allí descritos comprenden 25 paquetes, la mayoría de los cuales son parte de sistemas integrados de contabilidad. Esos paquetes permiten usar muchas características de controles de inventarios. En los casos más sencillos, como Microsoft Profit, sólo se encuentran las características básicas de control de inventarios, que incluyen informes de nivel de inventario e informes de existencias. La mayoría de los demás paquetes tienen características como métodos de costeo LIFO y FIFO, varias localidades, precios múltiples de artículos y análisis ABC de inventarios.

Algunos desarrolladores destinan sus productos sólo para mercados específicos. EFC Systems vende Kairos y Vanguard para aplicaciones agrícolas, principalmente. Otras empresas ofrecen varios productos para algunos segmentos del mercado. Uno de los mejores productos referidos lo ofrece Realworld Corporation, que vende Realworld Inventory Control e lnventory Plus. El primero puede manejar más de 1 000 bodegas, varias estructuras de precios y descuentos por cantidad. lnventory Plus agrega la serialización, control de pérdidas y codificación de barras. Este producto está diseñado para los detallistas, mientras que el primero es para ventas al mayoreo.

Muy pocos de los productos reseñados en este nivel de calidad permiten pronosticar, dimensionar lotes y calcular puntos de reorden. Una excepción es Integrated Accounting System, de CODA, Inc. Este es uno de los mejores sistemas, diseñado para computadoras DEC, IBM AS/400 y HP. Tiene una función de administración de inventario que admite varias bodegas y efectúa cálculos de punto de reorden y cantidades de pedido. De acuerdo con Tom Wittman, usuario del sistema en su empresa de pesca comercial, éste permite rastrear el estado de los inventarios en cinco localidades de manera simultánea.

La mayoría de los sistemas reseñados podrían ser adecuados para empresas pequeñas y medianas. Sin embargo, no incluyen las funciones que pueden requerir empresas grandes, como: 1) rastreo de los artículos en puntos de venta, 2) conexión de varias localidades en red, 3) una base central de datos que pueda ser consultada por localidades satélite, 4) implementación de intercambio electrónico de datos, 5) pronóstico de la demanda, 6) dimensionamiento de lote, y 7) cálculos de punto de reorden y de nivel de servicio. Los sistemas con estas propiedades son regularmente sistemas llave en mano.

Éstos son componentes y programas (hardware y sofware) integrados que se venden como unidad. Los fabricantes de computadoras centrales y estaciones de trabajo, como HP e IBM, y algunos vendedores especializados ofrecen esos sistemas. Uno de estos paquetes se analiza en la Aplicación instantánea del capítulo 5.

4.11 NOTAS HISTÓRICAS Y TEMAS ADICIONALES

El interés en el uso de modelos matemáticos para controlar el reabastecimiento de las existencias data de la primera parte del siglo XX. El desarrollo del modelo simple de cantidad económica de pedido se le suele atribuir a Ford Harris (1915). También se le reconoce a R. H. Wilson (1934) su análisis del modelo. El procedimiento sugerido para el modelo de descuento para todas las unidades parece deberse a Churchman, Ackoff y Arnoff (1957), y el modelo de descuento incremental a Hadley y Whitin (1963). Kenneth Arrow presenta una discusión excelente de los motivos económicos para mantener inventarios en el primer capítulo del libro de Arrow, Karlin y Scarf (1958).

También parece que Hadley y Whitin (1963) fueron los primeros en considerar problemas de restricción de presupuesto y espacio, aunque varios investigadores han estudiado después esos problemas. Rosenblatt (1981) deduce resultados parecidos a los de la sección 4.8 y considera varias extensiones que no se trataron en esa sección.

El interés académico en los problemas de administración de inventario creció súbitamente a finales de los años cincuenta y principios de los sesenta, cuando se publicaron varios textos,

Aplicación instantánea

RECONOCIMIENTO A MERVYN'S POR SU SISTEMA DE CONTROL DE INVENTARIOS

Mervyn's, una división de Dayton-Hudson Corporation, es detallista de alimentos secos con base en Hayward, California. En 1994 recibió un reconocimiento por su sistema computarizado de pronóstico y control de inventario (Donnelly, 1994).

Desde los últimos años de la década de los ochenta, la revista *Computerworld* y la Institución Smithsonian ha premiado organizaciones públicas y privadas por los usos eficaces de la tecnología de la información. Los nominados para el premio son seleccionados por un comité de 100 miembros, ejecutivos en el campo de tecnología de la información. Los ganadores se eligen con jurados pequeños de cada área. Las áreas en que se otorgan premios incluyen las finanzas, el gobierno, la manufactura, la medicina y los negocios.

Mervyn's fue seleccionada en 1994 como empresa ganadora de la categoría de Comercios y Servicios Afines por su sistema de administración de inventario en almacén planeado (*Planned Store Inventory, PSI*). Este complejo sistema coordina los datos de venta de 286 tiendas en todo el país. Según la gerencia de Mervyn's, con el PSI se redujeron los inventarios estimados en 60 millones de dólares, sin menoscabo de los niveles de servicio en las tiendas. En el corazón del PSI se encuentra un módulo de pronóstico que permite a los gerentes de tienda predecir la demanda de colores, tamaños y diseños específicos en su almacén. Una de las posibilidades únicas del sistema que se tiene en cuenta en el pronóstico es el lugar geográfico y la demografía del mercado, para diferenciar patrones de venta en almacenes ubicados en distintas partes del país.

Cada tienda tiene un promedio de 50 000 unidades de almacenamiento o números de parte individuales, así que coordinar los datos de las 286 tiendas es una tarea masiva de procesamiento de información. El sistema usa la última tecnología cliente-servidor, enlazando 486 computadoras personales a un multiprocesador basado en UNIX, servidor de base de datos.

El obstáculo más serio en la implementación del sistema PSI fueron los mismos gerentes de tienda. Los almacenes de Mervyn's se enorgullecen de suministrar altos niveles de servicio al cliente. Su objetivo es 95% de nivel de servicio, es decir, poder satisfacer la demanda del consumidor durante 95% del tiempo con existencias en almacén. Los gerentes de tienda estaban escépticos respecto a continuar suministrando esos niveles de servicio si se bajaban los niveles de inventario. De hecho, el sistema PSI recomendaba reducciones de inventarios en tienda hasta de 50% para algunos artículos.

El sistema de Mervyn's usa una combinación de técnicas de modelado matemático y tecnología cliente-servidor para guiar los niveles de inventario de las tiendas individuales y de artículos con ajustes continuos basados en observaciones de venta. Los inventarios se reducen aplicando un método justo a tiempo al reabastecimiento de las existencias. Según los competidores de Mervyn's, se contemplan los niveles generales de inventario y se ajustan de acuerdo con factores predeterminados.

Mervyn's es uno de los miembros fundadores del Instituto de Gerencia de Menudeo, de la Universidad Santa Clara, y siempre ha mantenido una robusta función de apoyo logístico. La dirección ha trabajado estrechamente relacionada con los profesores de la universidad en otros proyectos, para asegurar que sus métodos de pronóstico y control de inventarios usen la teoría y la tecnología de punta. El menudeo es un campo muy competitivo, y los buenos actores son los que reconocen que ya no son adecuadas las técnicas gerenciales donde las decisiones se toman a la fuerza. (Vea la Aplicación breve sobre Wal-Mart en el capítulo 1.) Wal-Mart es otro detallista que pone gran énfasis en la logística y las operaciones.)

además de los mencionados, incluyendo a Whitin (1957); Magee y Boodman (1967); Bowman y Fetter (1961); Fetter y Dalleck (1961); Hanssmann (1961); Start y Miller (1962); Wagner (1962), y Scarf, Gilford y Shelly (1963).

La gran mayoría de las investigaciones publicadas sobre sistemas de inventarios a partir de la década de los sesenta versa acerca de modelos estocásticos, que es el tema del siguiente capítulo. Más recientemente, Barbosa y Friedman (1978), y Schwarz y Schrage (1971) han considerado extensiones del modelo de CEP.

4.12 Resumen

Este capítulo presenta varios modelos muy usados para controlar inventarios cuando se conoce la demanda. Se describen los diversos *motivos económicos para mantener inventarios*, entre los que se encuentran las economías de escala, las incertidumbres, la especulación y la logística. Además se mencionan algunas de las *características físicas de los sistemas de inventario* importantes para determinar la complejidad y aplicabilidad de los modelos a los proble-

mas reales. Entre ellas figuran los supuestos sobre la demanda; el tiempo de demora; el tiempo de revisión; la acumulación de pedidos o la pérdida de ventas, y los cambios que se efectúan en el inventario con el paso del tiempo.

Hay tres clases importantes de costos en la administración de inventarios: *costos de mantener el inventario*; *costos de pedido o de preparación*, y *costo de penalización* por no cumplir con la demanda. El costo de mantener el inventario se suele expresar como el producto de una tasa de interés por el costo del artículo.

El abuelo de todos los modelos de control de inventario es el *modelo de la cantidad económica de pedido, o modelo CEP*, en el que se supone que la demanda es constante, no se permiten agotamientos y sólo hay costos de inventario y pedido. El tamaño óptimo de lote se calcula con la fórmula clásica de la raíz cuadrada. La primera extensión del modelo CEP descrito fue el caso en que los artículos se producían internamente con una tasa finita de producción. Se demostró que, en este caso, el tamaño óptimo de lote puede obtenerse mediante la fórmula de la CEP con un costo de inventario modificado.

Se describieron dos tipos de *descuentos por cantidad: en todas las unidades*, donde el precio descontado es válido para todas las unidades de un pedido, y en el *incremental*, donde el descuento sólo se aplica a las unidades adicionales por encima del punto límite. La solución óptima al caso de todas las unidades se presenta con frecuencia en un punto límite, mientras que la del caso incremental casi nunca se presenta en esos puntos de rompimiento.

En la mayor parte de los sistemas reales el gerente de inventarios no puede pasar por alto las interacciones existentes entre los productos. Esas interacciones imponen *restricciones* al sistema; pueden deberse a limitaciones en el espacio disponible para mantener el inventario o en el presupuesto disponible para comprar los artículos. Se han descrito ambos casos, demostrando que cuando la relación del valor del artículo o el espacio ocupado entre el costo de inventario es igual para todos los artículos, puede obtenerse con facilidad una solución al problema con restricciones. Cuando no se cumple esta condición, la formulación requiere introducir un *multiplicador de Lagrange.* El valor correcto de este multiplicador se puede precisar vía tanteos o mediante algún tipo de búsqueda. La presencia del multiplicador reduce los tamaños de lote aumentando los costos de mantenimiento del inventario.

El modelo de tasa de producción finita de la sección 4.6 fue ampliado a varios productos bajo el supuesto de que todos se fabrican en una sola máquina. Suponiendo que hay una *política de ciclo de rotación*, en la que se fabrica un lote de cada producto en cada ciclo, se demuestra cómo determinar el tiempo del ciclo que minimiza la suma de los costos anuales de preparación e inventario para todos los productos. Los ciclos de rotación permiten contar con una forma directa para programar la producción en una máquina, pero pueden ser subóptimos cuando las tasas de producción o demanda difieren mucho, o cuando los costos de preparación dependen de la secuencia.

El capítulo concluye con una presentación general de diversos sistemas comerciales de control de inventario. Existen muchos productos disponibles, cuyo precio va de menos de 100 dólares a decenas de miles de dólares. Muchos de los productos diseñados para computadoras personales son partes de sistemas integrados de contabilidad. Casi ninguno de esos productos contiene cálculos para pronósticos de artículos, dimensionamiento de lote y puntos de reorden.

Problemas adicionales sobre modelos deterministas de inventario

31. Peet's Coffees, en Menlo Park, California, vende filtros Melitta #101 para café a una tasa relativamente constante de 60 cajas mensuales. Esos filtros se piden a un proveedor en Trenton, Nueva Jersey. Al gerente de Peet's le interesa aplicar teoría de inventarios para calcular la mejor estrategia de reabastecimiento para los filtros.

    Peet's paga 2.80 dólares por caja de filtros, y estima que los costos fijos del tiempo de empleado para colocar y recibir pedidos asciende a 20 dólares, aproximadamente. Usa 22% de tasa de interés anual para calcular sus costos de inventario.

    a) ¿De qué tamaño debe ser el pedido de Peet's a su proveedor en Trenton, y con qué frecuencia debe hacer sus pedidos?

b) Suponga que se necesitan tres semanas para recibir un embarque. ¿Qué inventario de filtros debe tener a mano cuando hace un pedido?

c) ¿Cuáles son los costos anuales promedio fijos y de inventario asociados con esos filtros, suponiendo que se adopte una política óptima?

d) La tienda de Peet's en Menlo Park es bastante pequeña. ¿En qué puede afectar esto a la solución que recomendó en la parte *a*)?

32. Un supermercado local vende una marca popular de champú a una tasa constante de 380 botellas por mes. El costo de cada una en el supermercado es 0.45 dólares, y ha estimado que le cuesta 8.50 dólares hacer un pedido. Suponga que los costos de inventario se basan en una tasa de interés anual de 25%. No se permiten agotamientos en las existencias del champú.

    a) Calcule el tamaño óptimo de lote que debe pedir el supermercado y el tiempo entre los pedidos para este producto.

    b) Si el tiempo de demora del suministro es dos meses, calcule el punto de reorden con base en el inventario disponible.

    c) Si este artículo se vende en .99 dólares, ¿cuál es la ganancia anual (sin gastos administrativos ni mano de obra) que produce?

33. Diskup fabrica varios productos para computadoras personales. Los discos de alta densidad de 3.5 pulgadas se fabrican a una tasa de 1 800 por día, y se embarcan a una tasa de 800 diarios. Esos discos se producen en lotes. Cada disco cuesta 0.20 dólares a la empresa, y los costos de inventario se basan en la tasa de interés anual de 18%. No se permiten agotamientos. Cada corrida de producción de los discos requiere recalibrar el equipo, y se calcula que esto representa un costo de 180 dólares.

    a) Calcule el tamaño óptimo de cada corrida de producción y el tiempo entre ellas.

    b) ¿Qué fracción del tiempo dedica la empresa a producir discos de alta densidad de 3.5 pulgadas?

    c) ¿Cuál es la máxima inversión monetaria que tiene la empresa en esos discos?

34. Berry Computer está considerando mover algunas de sus operaciones trasatlánticas para reducir costos de trabajo. En Estados Unidos, producir el circuito principal le cuesta a Berry $75, mientras que en el extranjero sólo $65. Los costos de almacenamiento se basan en un interés de 20% anual, y por mucho tiempo la demanda se ha estabilizado en 200 unidades a la semana. Asumamos que la instalación cuesta $20 tanto en Estados Unidos como en el extranjero, y que los tiempos de producción óptimos son un mes, en Estados Unidos, y seis meses en el extranjero.

    a) Calcular el costo promedio anual de producción, almacenamiento e instalación en cada sitio, asumiendo que en cada caso se emplea una solución óptima. Con base en estos resultados, ¿qué lugar conviene más?

    b) Calcular el valor del conducto del inventario en cada caso. (El conducto del inventario es el inventario en orden.) ¿Comparar conductos de inventario afecta la conclusión a la que accedimos en *a*)?

    c) ¿Considerar otros factores puede favorecer las producciones local o trasatlántica?

35. Un gran fabricante de productos para el hogar compra externamente un glicérido que incorpora a uno de sus jabones desodorantes. Consume un promedio de 40 libras mensuales y lo hace en forma bastante estable. Para calcular los costos de inventario usa una tasa de

interés anual de 23%. Puede comprar ese producto a dos proveedores, A y B. A le ofrece el siguiente esquema de descuento para todas las unidades:

Tamaño de pedido	Precio por libra
$0 \leq Q < 500$	\$1.30
$500 \leq Q < 1{,}000$	1.20
$1{,}000 \leq Q$	1.10

mientras que B le ofrece el siguiente esquema de descuentos incrementales: 1.25 dólares por libra para pedidos menores o iguales a 700 libras, y 1.05 dólares por libra adicional a las 700 libras. Suponga que el costo de procesar el pedido en cada caso es 150 dólares. ¿Qué proveedor debe elegir?

36. Al presidente de Filtrapur le entusiasmó aplicar las cantidades económicas de pedido para planear los tamaños de sus corridas de producción, e instituyó el dimensionamiento de lotes basado en los valores de CEP antes de estimar bien sus costos. Para determinada línea de filtros, que tenía una demanda anual de 1 800 unidades por año, a un costo de 2.40 dólares por unidad, supuso un costo de inventario basado en 30% de interés anual, y un costo de preparación de 100 dólares. Tiempo después, cuando el departamento de contabilidad de costos pudo efectuar un análisis, encontró que el valor adecuado de la tasa de interés se acercaba más al 20%, y que el costo de preparación era de unos 40 dólares. ¿Cuáles fueron los costos adicionales de preparación e inventario anuales promedio en que se incurrió al usar los valores equivocados?



37. Considere el esquema de descuento por furgón que muestra la figura 4-13. Suponga también que $M = 500$ unidades, $C = 10$ dólares por unidad, y que un furgón completo de 500 unidades cuesta 3 000 dólares.

    a) Trace una gráfica del costo promedio por unidad, $C(Q)/Q$, suponiendo este esquema.

    b) Suponga que las unidades se consumen a una tasa de 800 por semana; que los costos de preparación y pedido son 2 500 dólares, y que los costos de inventario se basan en una tasa anual de interés de 22%. Grafique la función $G(Q) = \lambda C(Q)/Q + K\lambda/Q + I(C(Q)/Q)Q/2$ y determine el valor óptimo de Q. (Suponga que 1 año = 50 semanas.)

    c) Repita la parte *b*) con $\lambda = 1\,000$ por semana y $K = 1\,500$ dólares.

38. Pedro Gómez desea iniciar un negocio de emparedados para contar con ingresos extra. Sin embargo, sólo cuenta con 100 dólares para sus compras iniciales. Divide sus necesidades en tres clases: pan, carnes y quesos, y condimentos. Estima que podrá usar todos los productos que compra antes de que se echen a perder, de modo que lo perecedero no es asunto importante. Los parámetros de costo y demanda se muestran a continuación.

	Panes	Carnes y quesos	Condimentos
Demanda semanal	6 paquetes	12 paquetes	2 libras
Costo por unidad	\$0.85	\$3.50	\$1.25
Costo fijo de pedido	12	8	10

La selección de estos costos fijos se basa en que esos artículos se compran en distintos lugares de la localidad. Incluyen el costo del tiempo de Pedro para efectuar las compras. Suponga que los costos de inventario se basan en 25% de interés anual.

a) Calcule las cantidades óptimas que debe comprar Pedro de cada tipo de producto, de tal forma que no rebase su presupuesto.

b) Si pudiera comprar todos los artículos en un mismo lugar, ¿podría eso alterar su solución? ¿Por qué?

39. El Taller de Miguel, de servicio y reparación automotriz, consume filtros de aceite a una tasa bastante estable de 2 400 por año. Su dueño estima que el costo de su tiempo para elaborar y procesar un pedido es de unos 50 dólares. El proveedor necesita un mes para surtir los filtros, y cada uno le cuesta 5 dólares al taller. Éste usa 25% anual de interés para calcular su costo de mantener el inventario.

a) Calcule la cantidad óptima de filtros que debe comprar el Taller de Miguel, y el tiempo óptimo de espaciamiento de los pedidos.

b) Calcule el nivel de inventario disponible en el momento en que se debe hacer un pedido.

c) Suponiendo que el Taller de Miguel aplica una política óptima de inventario para los filtros de aceite, ¿cuál es el costo anual de inventario y pedido para este artículo?



40. Una empresa importadora y exportadora ha rentado 20 000 pies cúbicos de espacio de almacenamiento para guardar seis artículos que importa del Lejano Oriente. Los datos relevantes de los seis artículos que se almacenarán en la bodega son:

Artículo	Demanda anual (unidades)	Costo por unidad	Espacio por unidad (pies^3)
Videocaseteras	800	$200.00	12
TV de color, 13 pulgadas	1 600	150.00	18
Videocasetes (caja de 10)	8 000	30.00	3
Audiocasetes (caja de 10)	12 000	18.00	2
Estéreos compactos	400	250.00	24
Teléfonos	1 200	12.50	3

Los costos de preparación para cada producto ascienden a 2 000 dólares por pedido, y los de inventario se basan en una tasa de interés anual de 25%. Calcule las cantidades óptimas de pedido para esos seis artículos, de tal modo que el espacio de almacenamiento nunca se rebase. (Sugerencia: Use una celda para el multiplicador de Lagrange, y haga pruebas con distintos valores hasta que se satisfaga la restricción de volumen tanto como sea posible.)

41. Un fabricante de tarjetas de felicitación debe determinar el tamaño de corridas de producción para cierta línea popular de tarjetas. La demanda de éstas ha sido bastante estable (2 millones por año), y en la actualidad el impresor las produce en tamaños de lote de 50 000. El costo de preparar cada corrida de producción es de 400 dólares.

Suponga que el material para cada tarjeta cuesta 0.35 dólares, la mano de obra 0.15 dólares y el costo de distribución 0.05 dólares. El departamento de contabilidad de la empresa ha establecido una tasa de interés para representar el costo de oportunidad de inversiones alternativas y almacenamiento: equivale a 20% del valor de cada tarjeta.

a) ¿Cuál es el valor óptimo de la CEP para esta línea de tarjetas de felicitación?

b) Calcule el costo adicional anual que resulta por usar el tamaño incorrecto de lote de producción.

42. En el problema 41 suponga que la imprenta decide tener en cuenta que la tasa de producción de las tarjetas no es infinita. Calcule el tamaño óptimo de cada corrida de producción, suponiendo que se producen 75 000 tarjetas por semana.

43. Pastelería "R" hornea sus propios pasteles o *pays* en una instalación que cuenta con un horno grande al que le caben 100 pasteles. "R" los vende con una tasa bastante estable de 86 por mes. Hacer cada uno le cuesta 2 dólares. Antes de cada horneada hay que limpiar el horno, para lo cual se requiere una hora del tiempo de cuatro trabajadores, y a cada uno de ellos se le pagan 8.00 dólares por hora. Los costos de inventario se basan en la tasa de interés anual de 18%. Los pasteles pueden durar tres meses en el mostrador.

    a) ¿Cuántos pasteles deben hornearse en cada corrida de producción? ¿Cuál es el costo anual de preparación e inventario de los pasteles?

    b) El dueño de Pastelería "R" desea comprar un horno nuevo que requiere la mitad del tiempo para limpieza y el doble de capacidad que el anterior. ¿Cuál es la cantidad óptima de pasteles que debe hornearse cada vez en el nuevo horno?

    c) El costo neto del horno nuevo (vendiendo el anterior) es 350 dólares. ¿Cuántos años tardará el horno nuevo en "pagarse a sí mismo"?

44. El restaurante japonés Kirei-Hana Steak House, en San Francisco, consume 3 000 libras de sirloin por mes. Yama Hirai, el nuevo gerente, terminó recientemente su maestría en administración. Supo que la carne se reabastece usando un valor de CEP de 2 000 libras. Ese valor se calculó suponiendo una tasa de interés de 36% anual. Suponga que el costo actual de la carne para el restaurante es de 4 dólares por libra.

    a) ¿Cuál es el costo de preparación usado para calcular el valor de cantidad económica de pedido?

    El señor Hirai recibió una oferta de un tablajero mayorista que ofrece un descuento de 5% si compra la carne en lotes de 3 000 libras o más.

    b) ¿Debe aceptar la oferta del mayorista? En caso afirmativo, ¿cuánto puede ahorrar?

    c) Como el señor Hirai tiene problemas con el idioma, al parecer no entendió bien la oferta. En realidad, el 5% de descuento sólo se aplica a cantidades superiores a 3 000 libras. ¿Debe aceptar esta oferta? En caso afirmativo, ¿cuánto se ahorra en este caso?



45. Green Buttons de Rolla, Missouri, abastece todas las tiendas de New Jersey Fabrics con ocho estilos distintos de botones para camisas de vestir para caballero. La máquina de inyección de plástico sólo puede producir un estilo de botón cada vez, y requiere bastante tiempo para reconfigurarla y cambiar de estilo. Como Green's ha sido contratada para surtir cantidades fijas de botones durante los tres meses siguientes, su demanda puede considerarse fija y conocida. Los datos relevantes para este problema aparecen en la siguiente tabla:

Tipo de botón	Ventas anuales	Tasa de producción (unidades/día)	Tiempo de preparación (horas)	Costo variable
A	25 900	4 500	6	$0.003
B	42 000	5 500	4	0.002
C	14 400	3 300	8	0.008
D	46 000	3 200	4	0.002
E	12 500	1 800	3	0.010
F	75 000	3 900	6	0.005
G	30 000	2 900	1	0.004
H	18 900	1 200	3	0.007

Suponga 250 días laborables por año. El departamento de contabilidad de Green's ha establecido 18% de interés anual para costo de capital y 3% de interés anual para el espacio de almacenamiento. Los costos de preparación son 20 dólares por hora necesaria para reconfigurar el equipo antes de procesar un nuevo estilo. Suponga que la empresa adopta una política de ciclo de rotación para producir los botones.

a) ¿Cuál es el tiempo óptimo de ciclo de rotación?

b) ¿De qué tamaño deben ser los lotes?

c) ¿Cuál es el costo promedio anual de inventario y preparaciones en la solución óptima?

d) ¿Qué obligaciones contractuales podría tener Green's con New Jersey Fabrics que evitaran la implementación de la política determinada en las partes *a*) y *b*)? Más específicamente, si Green's convino enviar tres embarques por año para cada estilo de botón, ¿qué política de producción recomendaría usted?

Apéndice 4-A

Deducción matemática para sistemas de cantidad económica de pedido con restricciones para varios productos

Se examinará el sistema usual de CEP para varios productos con una restricción de presupuesto, descrito en la sección 4.8. El problema matemático consiste en determinar valores de las variables $Q_1, Q_2, \ldots, Q_n$ para

$$\text{Minimizar} \sum_{i=1}^{n} \left[\frac{h_i Q_i}{2} + \frac{K_i \lambda_i}{Q_i} \right]$$

sujeto a

$$\sum_{i=1}^{n} c_i Q_i \leq C.$$

Sean CEP_i los valores CEP respectivos sin restricción. Existen entonces dos posibilidades:

$$\sum_{i=1}^{n} c_i \text{CEP}_i \leq C, \tag{1}$$

$$\sum_{i=1}^{n} c_i \text{CEP}_i > C. \tag{2}$$

Si la ecuación (1) es válida, la solución óptima es la trivial, es decir, igualar $Q_i = \text{CEP}_i$. Si la ecuación (2) es válida, se tiene la garantía de que la restricción determina la solución óptima. Esto significa que la restricción puede escribirse como sigue:

$$\sum_{i=1}^{n} c_i Q_i = C.$$

En este caso se introduce el multiplicador de Lagrange θ, y ahora el problema consiste en determinar $Q_1, Q_2, \ldots, Q_n$, y θ que resuelvan el problema sin restricciones:

$$\text{Minimizar } G(Q_1, Q_2, \ldots, Q_n, \theta) = \sum_{i=1}^{n} \left(\frac{h_i Q_i}{2} + \frac{K_i \lambda_i}{Q_i} \right) + \theta \sum_{i=1}^{n} (c_i Q_i - C).$$

Las condiciones necesarias de optimalidad son

$$\frac{\partial G}{\partial Q_i} = 0 \qquad \text{para } i = 1, \ldots, n$$

y

$$\frac{\partial G}{\partial \theta} = 0.$$

Las primeras condiciones n producen

$$\frac{h_i}{2} - \frac{K_i \lambda_i}{Q_i^2} + \theta c_i = 0 \qquad \text{para } i = 1, \ldots, n.$$

Al reordenar los términos se obtiene

$$Q_i = \sqrt{\frac{2K_i \lambda_i}{h_i + 2\theta c_i}} \qquad \text{para } i = 1, \ldots, n,$$

y también se tiene la condición final

$$\sum_{i=1}^{n} c_i Q_i = C.$$

Considérese ahora el caso en que $c_i/h_i = c/h$, independiente de i. Si se dividen numerador y denominador entre h_i puede escribirse

$$Q_i = \sqrt{\frac{2K_i \lambda_i}{h_i}} \sqrt{\frac{1}{1 + 2\theta c/h}}$$

$$= \text{CEP}_i \sqrt{\frac{1}{1 + 2\theta c/h}}$$

$$= \text{CEP}_i m,$$

donde

$$m = \sqrt{\frac{1}{1 + 2\theta c/h}}.$$

Al sustituir el resultado de Q_i en la restricción se obtiene

$$\sum_{i=1}^{n} c_i \text{CEP}_i m = C$$

o

$$m = C \bigg/ \left(\sum_{i=1}^{n} c_i \text{CEP}_i \right).$$

Apéndice 4-B

Glosario de notación para el capítulo 4

c = Costo proporcional al pedido.

CEP = Cantidad económica de pedido (tamaño de lote).

$G(Q)$ = Costo anual promedio asociado con el tamaño de lote Q.

h = Costo de mantener el inventario por unidad de tiempo.

h' = Costo modificado de mantener el inventario, para el modelo de tasa finita de producción.

I = Tasa anual de interés usada para calcular el costo de mantener el inventario.

K = Costo de preparación o costo fijo de pedido.

λ = Tasa de demanda (unidades por unidad de tiempo).

P = Tasa de producción para el modelo de tasa finita de producción.

Q = Tamaño de lote o tamaño del pedido.

s_i = Tiempo de preparación para el producto i (véase la sección 4.9).

T = Tiempo de ciclo; tiempo entre colocación de pedidos sucesivos.

τ = Tiempo de demora del pedido.

θ = Multiplicador de Lagrange para el modelo con restricción de espacio (véase la sección 4.8).

w_i = Espacio ocupado por una unidad del producto i (véase la sección 4.8).

W = Espacio total disponible (véase la sección 4.8).

Bibliografía

Arrow, K. A., S. Karlin y H. Scarf., Eds. *Studies in the Mathematical Theory of lnventory Production.* Stanford, CA: Stanford University Press, 1958.

Barbosa, L. C. y M. Friedman. "Deterministic Inventory Lot Size Models—A General Root Law". *Management Science* 23 (1978), pp. 820-829.

Bowman, E. H. y R. B. Fetter. *Analysis for Production Management.* Homewood, IL: Richard D. Irwin, 1961.

Churchman, C. W., R. L. Ackoff y E. L. Arnoff. *lntroduction to Operations Research.* Nueva York: John Wiley & Sons, 1957.

Cohn, M. "Keeping Track of fue Pieces". *Accounting Technology* 10, núm. 2, (1994), pp. 18-37.

Donnelly, H. "Technology Awards: Recognizing Retailing's Best". *Stores* 76, núm. 10, (1994), pp. 52-56.

Fetter, R. B. y W. C. Dalleck. *Decision Models for lnventory Management.* Homewood, IL: Richard D. Irwin, 1961.

Hadley, G. J. y T. M. Whitin. *Analysis of lnventory Systems,* Englewood Cliffs, NJ: Prentice Hall, 1963.

Hanssmann, F. "A Survery of Inventory Theory from the Operations Research Viewpoint". En *Progress in Operations Research* 1, ed. R. L. Ackoff. Nueva York: John Wiley & Sons, 1961.

Harris, F. W. Operations and Cost (Factory Management Series). Chicago: Shaw, 1915.

Johnson, L. A. y D. C. Montgomery. *Operations Research in Production Managing, Scheduling and Inventory Control*. Nueva York: John Wiley & Sons, 1974.

Magee, J. F. y D. M. Boodman. *Production Plalzning and lnventory Control.* 2a. Ed., Nueva York: McGraw Hill, 1967.

Maxwell, W. L. "The Scheduling of Economic Lot Sizes". *Naval Research Logistics Quarterly* 11 (1964), pp. 89-124.

Plossl, G. W. y O. W. Wight. *Production and lnventory Control.* Englewood Cliffs, NJ: Prentice Hall, 1967.

Rosenblatt, M. J. "Multi-Item Inventory System with Budgetary Constraint: A. Comparison between fue Lagrangian and the Fixed Cycle Approach". *lnternational Journal of Production Research.* 19 (1981), pp. 331-339.

Scarf, H. E., D. M. Gilford y M. W. Shelly. *Multistage lnventory Models and Techniques.* Stanford CA: Stanford University Press, 1963.

Schwarz, L. B. y L. Schrage. "Optimal and Systems Policies for Multiechelon Production/Inventory Assembly Systems". *Management Science* 21 (1971), pp. 1285-1294.

Starr, M. K. y D. W. Miller. *lnventory Control: Theory and Practice.* Englewood Cliffs, NJ: Prentice Hall, 1962.

Wagner, H. M. *Statistical Management of Inventory Systems.* Nueva York: John Wiley & Sons,1962.

Whitin, T. M. *The Theory of Inventory Management.* Edición revisada. Princeton University Press, 1957.

Wilson, R. H. "A Scientific Routine for Stock Control". *Harvard Business Review* 13 (1934), pp. 116-128.

Capítulo **Cinco**

Control de inventarios sujetos a demanda incierta

Panorama general del capítulo

Propósito

Comprender cómo debe manejarse la incertidumbre (aleatoriedad) en la demanda al momento de calcular las políticas de reabastecimiento para un solo artículo en existencia.

Puntos clave

1. *¿Qué es la incertidumbre y cuándo debe suponerse su existencia?* Hablamos de incertidumbre cuando la demanda es una variable aleatoria, y ésta se define según la distribución de probabilidad, que por lo general se calcula a partir de un historial anterior de demandas. En la práctica, es común suponer que la demanda sigue una distribución normal. Cuando la demanda se supone normal, sólo se necesita calcular la media, μ, y la varianza, σ^2. Claramente, la demanda es incierta hasta cierto grado en todas las aplicaciones del mundo real. ¿Qué valor tiene entonces el análisis de los capítulos 3 y 4, cuando la demanda se asumió como conocida? El capítulo 3 se enfocó en cambios sistemáticos o predecibles en el patrón de demanda: crestas y valles. En el capítulo 4 pudimos percatarnos de que los artículos únicos son útiles si la variación de la demanda es baja en relación con la media. Ahora estudiaremos objetos cuya principal variación se debe a la incertidumbre más que a causas predecibles.

   Si la demanda se describe por una variable aleatoria, no queda claro cuál debe ser el criterio de optimización, debido a que la función de costo es también una variable de ese tipo. Para manejar esto, suponemos que el objetivo es minimizar los costos esperados. El uso del operador de expectativas se justifica por la ley de los grandes números proveniente de la probabilidad, ya que el problema de control de inventarios abarca muchos periodos de planeación. La ley de los grandes números garantiza que el promedio aritmético de los costos incurridos y los costos esperados se incremente conforme aumenta el número de periodos de planeación.

2. *El modelo del repartidor de periódicos.* Un proveedor de periódicos decide cuántos comprar para vender durante el día. Como la demanda diaria es muy variable, se ha modelado con una variable aleatoria, *D*. Supongamos que *Q* es el número de diarios que compra. Si *Q* es demasiado grande, se queda con diarios sin vender, y si *Q* es demasiado pequeño, no satisface algunos pedidos. Si c_O es el costo unitario de excedente de

mercancías, y c_u es el costo por faltantes de mercancía; entonces, mostramos que el número óptimo de diarios que debe comprar al iniciar el día, digamos Q^*, satisface:

$$F(Q^*) = c_u/(c_u + c_o)$$

donde $F(Q^*)$ es la función de distribución acumulada de D evaluada en Q^* (que es la misma probabilidad de que la demanda sea menor o igual a Q^*).

3. *Sistemas de punto de reorden-tamaño del lote.* Es apropiado utilizar el modelo de repartidor de periódicos para un problema que esencialmente reinicia desde cero cada periodo. El periódico de ayer ya no tiene valor en el mercado, salvo por el posible valor que pudiera tener el papel en sí. Sin embargo, la mayoría de las situaciones de control de inventarios que se encuentran en el mundo real no son como esta. Los artículos que no se venden continúan teniendo valor en el mercado durante muchos periodos. Para estos casos usamos un enfoque que es esencialmente una extensión del modelo CEP del capítulo 4.

   El sistema de punto de reorden-tamaño del lote se basa en la suposición de que los inventarios se revisan en forma continua más que periódica. Es decir, que el estado del sistema se conoce en todo momento. Éste consta de dos variables de decisión: Q y R. Q es el tamaño de pedido y R es el punto de reorden. Así, cuando el inventario de existencias alcanza R, se hace un pedido de Q unidades. El modelo también permite un tiempo positivo de demora de pedido τ. La demanda sobre el tiempo de demora es la clave para la incertidumbre en el problema, ya que el tiempo de demora es el tiempo de respuesta del sistema. Supongamos que D representa la demanda sobre el tiempo de demora, y digamos que $F(t)$ es la función de distribución de demanda acumulada de D. Los parámetros de costo incluyen un costo fijo de pedido K, un costo de penalización unitario para demanda insatisfecha p, y un costo unitario de tiempo de retención h. Interpretemos λ como la tasa anual promedio de la demanda (la esperada en un año). Más adelante en esta sección se muestra que los valores óptimos de Q y R satisfacen las siguientes dos ecuaciones no lineales simultáneas:

$$Q = \sqrt{\frac{2\lambda[K + pn(R)]}{h}}$$

$$1 - F(R) = Qh/p\lambda.$$

   Para resolverlas se requiere un método de solución iterativa de ida y vuelta. Proporcionamos detalles del método sólo cuando la distribución de la demanda del tiempo de demora es normal. La convergencia por lo general se presenta rápidamente. Una aproximación rápida pero burda sería establecer Q = CEP y despejar R de la segunda ecuación. En la mayoría de los casos esto da buenos resultados.

4. *Niveles de servicio en los sistemas (Q, R).* Contamos con dos tipos de servicio: el Tipo 1 es la probabilidad de no quedarse sin mercancía en el tiempo de demora y está representado por el símbolo α. El Tipo 2 es la proporción de las demandas que se surten con las existencias (también conocido como índice de satisfacción) y se representa mediante el símbolo β. Es muy sencillo encontrar el óptimo (Q, R) sujeto al objetivo de Tipo 1. Simplemente se halla R desde $F(R) = \alpha$ estableciendo Q = CEP. Desafortunadamente, con el término "servicio" casi siempre se hace referencia al Tipo 2, y en ese caso encontrar (Q, R) es más difícil. Para el servicio Tipo 2, consideramos solamente la distribución normal. La solución requiere utilizar las tablas de pérdidas estandarizadas $L(z)$ que aparecen al final de este libro. Al igual que sucede con el modelo de costo, establecer Q = CEP y despejar R por lo general da buenos resultados cuando se desea evitar la molestia de un procedimiento iterativo.

En este capítulo se abordan la relación entre el control de inventarios y el pronóstico, y la forma en que normalmente se actualizan los cálculos de la desviación media y estándar usando el suavizamiento exponencial. La sección concluye con un análisis acerca de la variabilidad del tiempo de demora y la forma en que se considera la incertidumbre adicional.

5. *Sistemas de revisión periódica bajo incertidumbre.* El modelo del repartidor de periódicos trata sobre productos que perecen rápidamente (después de un periodo). Sin embargo, los modelos de revisión periódica también son pertinentes cuando un producto no vendido puede utilizarse en periodos futuros. En este caso, la forma de la política óptima se conoce como una política (s, S). Digamos que u es el inventario inicial en cualquier periodo. Entonces la política (s, S) es

   Si $u \leq s$, haga el pedido para S (esto es, pida $S - u$).

   Si $u > s$, no pida nada.

   Desgraciadamente, es mucho más complicado hallar los valores óptimos de (s, S) que encontrar la política (Q, R) óptima y se encuentra más allá del ámbito de este libro. También analizaremos brevemente los niveles de servicio en sistemas periódicos de revisión.

6. *Sistemas multiproducto.* Prácticamente todos los problemas de control de inventarios que ocurren en el contexto de planeación de operaciones involucran múltiples productos. Una cuestión que surge en los sistemas multiproducto es determinar la cantidad de esfuerzo que una persona debe emplear para manejar cada artículo. Es evidente que algunos artículos son más valiosos para el negocio que otros. El sistema ABC es un medio para clasificar artículos. Los artículos se disponen en secuencia de orden decreciente de acuerdo con el volumen anual de ventas o uso en dólares. Al ordenar los artículos de esta manera y graficar el volumen acumulado en dólares se obtiene una curva de crecimiento exponencial denominada curva de Pareto. Por lo general, 20 por ciento de los artículos representan el 80 por ciento del volumen anual en dólares (artículos A), el siguiente 30 por ciento de los artículos representa normalmente el posterior 15 por ciento del volumen en dólares (artículos B), y el 50 por ciento restante de los artículos representa el 5 por ciento final del volumen en dólares (artículos C). Los artículos A deben recibir mayor atención que los demás. Sus niveles de inventario necesitan revisarse con frecuencia y se requiere llevar un alto nivel de servicio. Los artículos B no necesitan un escrutinio tan cercano, y los artículos C, por lo general, no se piden en grandes cantidades.

7. *Otros problemas.* El estudio de los modelos de inventario estocásticos que se lleva a cabo en este capítulo revela apenas la punta del iceberg en términos de la gran cantidad de investigación realizada sobre este tema. Dos importantes áreas de investigación son los sistemas de inventario multinivel, y los sistemas de inventario perecederos. Un sistema de inventario multinivel es aquel en el que se almacenan los artículos en ubicaciones múltiples unidas por una red. Las cadenas de abastecimiento, que se comentarán con mayor detalle en el capítulo 6, son un sistema de ese tipo. Otra área importante de investigación son los artículos que cambian durante el almacenamiento, lo que afecta su tiempo de vida útil. Una clase de dichos artículos son los productos perecederos, que tienen un tiempo de vida fijo conocido con antelación, e incluyen alimentos, productos farmacéuticos y películas fotográficas. Un problema relacionado con ello sería manejar artículos sujetos a su obsolescencia. La obsolescencia difiere de la corruptibilidad en que el tiempo de vida útil de un artículo sujeto a la obsolescencia no puede predecirse con antelación. Los modelos matemáticos para analizar dichos problemas son muy complejos y están fuera del ámbito de esta obra.

   En la administración de inventarios surge una cuestión práctica concerniente a la forma más eficiente de intercambiar información entre ubicaciones múltiples. Un método de ese tipo se conoce como intercambio electrónico de datos (IED), el cual ha ganado popularidad en las aplicaciones detallistas. IED es un término general para los enlaces electrónicos entre bodegas y almacenes, fábricas y bodegas, proveedores y compradores, etc. Existen varios protocolos IED disponibles, incluyendo algunos sistemas basados en red.

En el éxito de cualquier empresa el manejo de la incertidumbre desempeña un papel importante. ¿Cuáles son las fuentes de incertidumbre que afectan a un negocio? En una lista parcial se deben mencionar la incertidumbre con respecto a las preferencias del consumidor y las tendencias en el mercado; a la disponibilidad y el costo de la mano de obra y los recursos; a los tiempos de suministro por parte de los proveedores; al clima y sus ramificaciones en el ámbito de la logística de las operaciones; la incertidumbre relativa a las variables financieras como precios de las acciones y tasas de interés, y a la demanda de productos y servicios.

Antes de la terrible tragedia del 11 de septiembre de 2001, muchos de nosotros solíamos llegar al aeropuerto 30 o 40 minutos antes de la hora de salida del vuelo. Ahora debemos estar dos horas antes. El aumento de la seguridad en los aeropuertos ha incrementado no sólo el promedio de tiempo estipulado, sino también la incertidumbre en cuanto a éste. Para compensar dicha incertidumbre, llegamos mucho antes de la hora programada para asegurar un margen mayor de tiempo. Este mismo principio se aplica al manejo de inventarios.

La incertidumbre de la demanda y su efecto sobre las estrategias de administración de inventarios son los temas de este capítulo. Por ejemplo, una tienda departamental no puede predecir con exactitud las ventas de cierto artículo en determinado día. Una aerolínea no puede pronosticar con precisión la cantidad de personas que decidirán viajar en determinado vuelo. Entonces, ¿cómo pueden tales empresas decidir la cantidad de artículos que conservarán en inventario, o elegir la cantidad de vuelos que habrán de programarse en una ruta dada?

Aunque no se pueden predecir las ventas exactas de un artículo o las cantidades de asientos ocupados en un avión, la experiencia puede proporcionar información útil para la planeación. Como se indicará en la sección siguiente, las observaciones previas de cualquier fenómeno aleatorio se pueden usar para estimar su *distribución de probabilidades.* Al cuantificar en forma adecuada las consecuencias de las decisiones incorrectas, un modelo matemático bien pensado del sistema que se estudia generará estrategias inteligentes. El objetivo casi siempre consiste en minimizar el costo esperado o maximizar las ganancias esperadas cuando existe incertidumbre.

Cuando comenzó a manifestarse la recesión económica de principios de los años noventa, algunas empresas que dependían del gasto directo del consumidor, como los de ventas al menudeo, padecieron graves pérdidas. Tanto Sears como Macy's, muy exitosas en el mercado estadounidense del menudeo, anunciaron mala operación en 1991; de hecho, Macy's estaba camino a la quiebra. Sin embargo, varios minoristas lograron éxitos resonantes durante esos tiempos de dificultad económica. En el giro de la moda, tanto The Gap, con sede en San Francisco, y The Limited, en Columbus, Ohio, se desempeñaron muy bien. Las tiendas Wal-Mart continuaron su ascenso, sobrepasando a Sears como el mayor minorista en Estados Unidos. La incertidumbre en la demanda, especialmente en la industria del vestido, es un factor importante. La administración inteligente de inventarios, frente a la incertidumbre, ciertamente desempeñó un papel clave en el éxito de las firmas mencionadas.

Aparentemente casi todos los problemas prácticos de administración de inventarios están caracterizados por cierto grado de incertidumbre en la demanda, por lo que podría cuestionarse el valor de los modelos deterministas de control de inventarios descritos en el capítulo 4. Hay dos razones para estudiar los modelos deterministas. Una es que forman una base para comprender los compromisos fundamentales que se encuentran en la administración de inventarios; otra es que, dependiendo del grado de incertidumbre en la demanda, pueden ser buenos modelos del sistema.

Para comprender mejor el segundo punto de vista, sea D la demanda de un artículo durante cierto intervalo de tiempo. Expresaremos a D como la suma de dos partes, D_{Det} y D_{Ale}. Esto es,

$$D = D_{Det} + D_{Ale},$$

en donde

$$D_{Det} = \text{Componente determinista de la demanda,}$$
$$D_{Ale} = \text{Componente aleatorio de la demanda.}$$

Existen varias circunstancias bajo las cuales es adecuado considerar que D es determinista, aunque D_{Ale} no sea cero. Algunas de ellas son:

1. Cuando la varianza del componente aleatorio, D_{Ale} es pequeña en relación con la magnitud de D.
2. Cuando la variación predecible es más importante que la variación aleatoria.
3. Cuando la estructura del problema es demasiado compleja para incluir una representación explícita de la aleatoriedad en el modelo.

Se presenta un ejemplo del caso 2 en el problema de planeación agregada. Aunque el error pronosticado de las demandas agregadas durante el horizonte de planeación pueda no ser cero, es más importante planear los cambios anticipados en la demanda que los no anticipados. Un ejemplo del caso 3 se presenta en la planeación de requerimientos de materiales (que se describirá con detalle en el capítulo 7). Las encrucijadas de las relaciones entre los diversos niveles de componente y artículos finales dificultan la incorporación de la incertidumbre de la demanda en el análisis.

Sin embargo, el componente aleatorio de la demanda para muchos artículos es demasiado importante como para pasarlo por alto. Mientras la demanda esperada por unidad de tiempo sea relativamente constante, y la estructura del problema no sea demasiado compleja, es preferible manejar explícitamente la incertidumbre de la demanda. Algunos de los modelos estocásticos de inventario más relevantes, y los aspectos clave que rodean a la incertidumbre, se examinan en este capítulo.[1]

Panorama general de los modelos descritos en este capítulo

Los modelos de control de inventarios sujetos a incertidumbre son básicamente de dos tipos: 1) de **revisión periódica** y 2) de **revisión continua**. (Recuerde la descripción dada al iniciar el capítulo 4. *Revisión periódica* significa que el nivel de inventarios se conoce sólo en puntos discretos del tiempo, y *revisión continua* quiere decir que dicho nivel se conoce siempre.) Los modelos de revisión periódica pueden ser para uno o varios periodos de planeación. Para los modelos de un periodo, el objetivo es balancear correctamente los costos de sobrantes (por pedir demasiado) y los de desabasto (por pedir muy poco). Los modelos de un periodo son útiles en varios contextos: planeación para tamaños iniciales de embarque con artículos de alta costura, políticas de pedido para productos alimenticios de corta duración, o determinar tamaños de corrida de producción para artículos con vida útil breve, como los periódicos. A causa de esta última aplicación, el modelo estocástico de inventarios para un periodo se ha llamado **modelo del repartidor de periódicos**.[2] Este modelo será el primero en ser descrito aquí.

La gran mayoría de los sistemas de control de inventarios basados en computadora que hay en el mercado, según la experiencia del autor, emplean alguna variante de los modelos de revisión continua que se describen después en este capítulo. En cierto sentido, son extensiones del modelo de cantidad económica de pedido que incorporan incertidumbre. En la práctica, su popularidad puede atribuirse a varios factores. Primero, que las políticas son fáciles de calcular e implementar. Segundo, que los modelos describen fielmente la mayoría de los sistemas en los que existe reabasto dinámico de los artículos de inventario bajo incertidumbre. Se incluye también una descripción detallada de los niveles de servicio. Los métodos de nivel de servicio se implementan con mayor frecuencia que los costos de penalización, porque en la práctica es difícil estimar estos últimos.

Los modelos estocásticos de inventario para varios periodos dominan las publicaciones profesionales sobre teoría de inventarios.

---

[1] Para quienes no estén familiarizados con el término *estocástico*, sólo es un sinónimo de *aleatorio*.

[2] Espero que nadie se ofenda por esta etiqueta. Algunos autores han usado los términos *voceador* o *vendedor de diarios* para este modelo, pero elegí conservar la nomenclatura original para una mayor fidelidad histórica.

En esta apasionante área de investigación se han obtenido resultados suficientes como para compilar un volumen especializado. Sin embargo, el objetivo de este libro es enfocar la metodología aplicada en el mundo real. Los modelos estocásticos para varios periodos proporcionan una buena perspectiva, aunque rara vez se implementan. Por otra parte, el nivel de complejidad matemática que requieren escapa a los alcances de esta obra. Por estas dos razones, aquí no se describen los modelos de inventario para varios periodos.

5.1 NATURALEZA DE LA ALEATORIEDAD

Con la intención de aclarar el significado de los términos *aleatoriedad* e *incertidumbre* en el contexto del control de inventarios, plantearemos un ejemplo.

Ejemplo *5.1*

Mac, dueño de un puesto de periódicos, los domingos compra varios ejemplares de *The Computer Journal*, semanario bastante conocido. Paga 25 centavos por cada ejemplar, que vende a 75 centavos. Los ejemplares que no vende durante la semana puede regresarlos a su proveedor, quien le paga 10 centavos por cada uno. El abastecedor puede reciclar el papel para imprimir los números futuros. Mac ha registrado la demanda del *Journal* cada semana. (Incluye la cantidad de ejemplares realmente vendida más la cantidad de pedidos que no pudo surtir.) Las demandas observadas durante cada una de las últimas 52 semanas fueron

15	19	9	12	9	22	4	7	8	11
14	11	6	11	9	18	10	0	14	12
8	9	5	4	4	17	18	14	15	8
6	7	12	15	15	19	9	10	9	16
8	11	11	18	15	17	19	14	14	17
13	12								

No hay un patrón notable en estos datos, de modo que es difícil predecir la demanda del *Journal* en cualquier semana dada. Sin embargo, la historia de la demanda de este artículo puede representarse en un histograma de frecuencias, que expresa la cantidad de veces que se observó cómo ocurría la demanda semanal durante el año. El histograma para este patrón de la demanda se ilustra en la figura 5-1.

El histograma de frecuencias se usa para estimar la probabilidad de que la cantidad de ejemplares del *Journal* que se venden cualquier semana sea un valor específico. Esos estimados de probabilidad se obtienen dividiendo entre 52 la cantidad de veces que se observó cada ocurrencia de la demanda durante el año. Por ejemplo, la probabilidad de que la demanda sea 10 se estima como $2/52 = .0385$, y

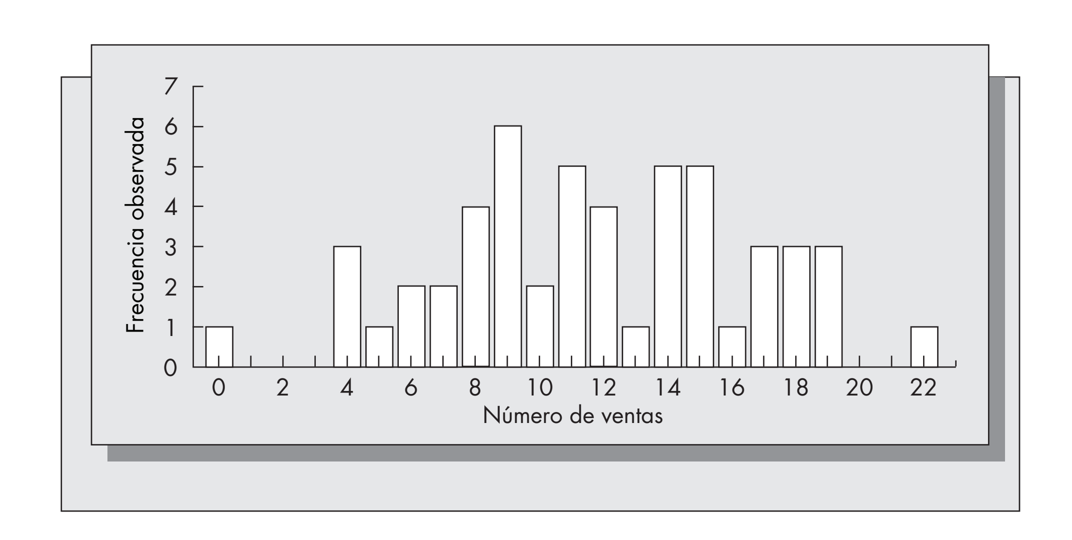


FIGURA 5-1 Histograma de frecuencias para una historia de 52 semanas de ventas de *The Computer Journal* en el puesto de Mac.

la probabilidad de que sea 15 es $5/52 = .0962$. El conjunto de todas las probabilidades se denomina *distribución empírica de probabilidades*. Las probabilidades acumuladas se pueden estimar también en forma parecida. Por ejemplo, la probabilidad de que se vendan nueve ejemplares o menos del *Journal* en alguna semana es $(1 + 0 + 0 + 0 + 3 + 1 + 2 + 2 + 4 + 6) = 19/52 = .3654$.

En análisis posteriores se pueden emplear las probabilidades empíricas, pero esto no es conveniente por varias razones. En primer lugar, es necesario mantener un registro de la historia de la demanda para cada artículo, lo que puede ser caro y tedioso. En segundo lugar, la distribución debe expresarse (en este caso) como 23 probabilidades distintas. Hay otros artículos que pueden tener un intervalo aún mayor de valores históricos. Por último, es más difícil calcular las políticas óptimas de inventarios con distribuciones empíricas.

La historia de la demanda suele aproximarse usando una distribución continua por dichas razones. La forma de la distribución elegida depende de la historia de la demanda y de lo fácil que sea emplearla. Con mucho, la distribución más usada para aplicaciones de inventarios es la normal. Una razón es la frecuencia con que parece modelar fielmente las fluctuaciones de la demanda. Otra es su comodidad. Sin embargo, el modelo normal de la demanda debe manejarse con cuidado, porque admite la posibilidad de valores negativos. Al usar la distribución normal para describir un fenómeno no negativo como la demanda, la posibilidad de una observación negativa debe ser suficientemente pequeña (menor que .01 bastaría para la mayoría de las aplicaciones) como para que no sea relevante.

Una distribución normal se determina mediante dos parámetros: la media μ y la varianza σ^2. Ambos pueden estimarse a partir de una historia de la demanda, la media de la muestra, $\overline{D}$, y la varianza de la muestra s^2. Sean $D_1, D_2, \ldots, D_n$ observaciones n de la demanda en el pasado. Entonces

$$\overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i,$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (D_i - \overline{D})^2.$$

Para los datos que se aprecian en la figura 5-1 se obtiene

$$\overline{D} = 11.73,$$

$$s = 4.74.$$

La función de densidad normal, $f(x)$, se expresa con la fórmula

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \qquad \text{para } -\infty < x < +\infty.$$

En las aplicaciones, $\overline{D}$ se sustituye por el estimador μ y s por el estimador σ.

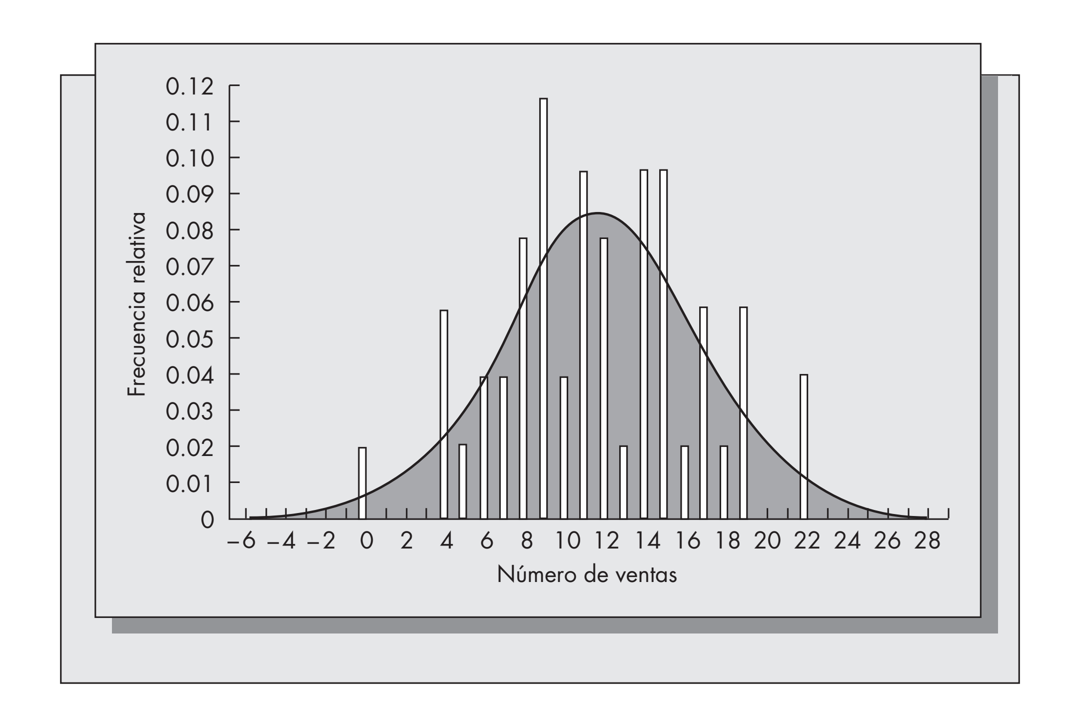
El *histograma de frecuencia relativa* es igual al histograma de frecuencias que se ilustra en la figura 5-1, con la excepción de que los elementos del eje y se dividen entre 52. La figura 5-2 muestra la función de densidad normal que resulta al hacer las sustituciones anteriores, sobrepuesto al histograma de frecuencia relativa.

El suavizamiento exponencial es utilizado en la práctica para actualizar recursivamente los estimados de la media y la desviación estándar de la demanda. La desviación estándar se estima con la desviación absoluta media (DAM). Tanto el suavizamiento exponencial como la desviación absoluta media se describieron con detalle en el capítulo 2. Sea $\overline{D}_t$ el estimado de la media después de observar la demanda D_t, y sea DAM_t el estimado de la DAM. Entonces

$$\overline{D}_t = \alpha D_t + (1 - \alpha)\overline{D}_{t-1},$$

$$\text{DAM}_t = \alpha\left|D_t - \overline{D}_{t-1}\right| + (1 - \alpha)\text{DAM}_{t-1},$$

FIGURA 5-2
Histograma
de frecuencia
y aproximación
normal.



en la que $0 < \alpha < 1$ es la constante de suavizamiento. Para la demanda con distribución normal,

$$\sigma \approx 1.25 * \text{DAM}.$$

Por lo regular se usa una constante de suavizamiento $\alpha \approx .1$ para asegurar la estabilidad en los estimados (véase en el capítulo 2 la selección de la constante de suavizamiento en detalle).

5.2 CRITERIO DE OPTIMIZACIÓN

La optimización en los problemas de producción generalmente implica determinar una regla de control mediante la cual se logre el costo mínimo. Sin embargo, cuando la demanda es aleatoria, el costo mismo en que se incurre también lo es, y ya no es obvio cuál debe ser el criterio de optimización. Virtualmente, el objetivo de todas las técnicas estocásticas de optimización aplicadas al control de inventarios es minimizar los costos *esperados.*

La razón para optar por el criterio del valor esperado es que los problemas de control de inventarios son problemas dinámicos, por lo general. Las decisiones se toman en forma repetitiva. La ley de los grandes números en la teoría de probabilidad establece que el promedio aritmético de muchas observaciones de una variable aleatoria converge hacia el valor esperado de la misma. Si en el contexto del problema de inventarios se adopta una regla de control que minimice lo que se estima gastar, entonces el promedio aritmético de los costos reales incurridos durante muchos periodos también será un mínimo.

En cierto contexto podrá suceder que el valor esperado no sea el mejor criterio de optimización. Cuando se adquiere un producto sólo una vez, no es obvio que lo adecuado sea minimizar los costos esperados. En ese caso, por lo general lo más adecuado es maximizar la

probabilidad de algún evento (como satisfacer una proporción de la demanda). Sin embargo, debido a la naturaleza dinámica de la mayoría de los problemas de producción, el criterio del valor esperado se usa virtualmente en todas las aplicaciones estocásticas de control de inventarios.

Problemas para las secciones 5.1 y 5.2

1. Suponga que Mac sólo registró la cantidad de revistas que vendió. ¿Sería esto una representación fiel de la demanda de *Journal*? ¿Bajo qué circunstancias se aproximarían la demanda real y la cantidad vendida, y bajo cuáles diferirían apreciablemente?
2. ¿Cuál es la diferencia entre las variaciones determinista y aleatoria en el patrón de las demandas? Cite un ejemplo de un problema real en el que sea importante la variación predecible, y otro en el que sea importante la variación aleatoria.
3. Oakdale Furniture Company usa un pegamento especial para ensamblar sus productos. Durante las últimas 36 semanas usó las siguientes cantidades de pegamento (en galones):

25	38	26	31	21	46	29	19	35	39	24	21
17	42	46	19	50	40	43	34	31	51	36	32
18	29	22	21	24	39	46	31	33	34	30	30

a) Calcule la media y la desviación estándar de esta muestra.

b) Considere los siguientes intervalos de clase para la cantidad de galones usados cada semana:

Menos de 20.
20-27.
28-33.
34-37.
38-43.
Más de 43.

Calcule la proporción de los datos que caen en cada uno de esos intervalos. Compare esas proporciones con las probabilidades de que una variable normal, con la media y la desviación estándar que calculó en la parte *a*, quede en cada uno de esos intervalos. Con base en la comparación de las proporciones observadas y las calculadas, suponiendo una distribución normal, ¿concluiría usted que la distribución normal se ajusta adecuadamente a esos datos? (Este procedimiento es en esencia el mismo que una prueba chi-cuadrada de bondad de ajuste.)

c) Suponga que la cantidad de galones de pegamento que se usan cada semana son variables aleatorias independientes, que tienen la distribución normal con la media y la desviación estándar calculadas en la parte *a*. ¿Cuál es la probabilidad de que la cantidad total de galones a usarse en seis semanas no sobrepase los 200 galones? (Sugerencia: La media de una suma de variables aleatorias es la suma de las medias, y la *varianza* de una suma de variables aleatorias *independientes* es la suma de las varianzas.)

4. En el problema 3, ¿qué otras distribuciones de probabilidad podrían describir con exactitud el uso de pegamento por parte de Oakdale?
5. Más que mantener registro de cada observación de la demanda, Betty Sucasas, miembro del personal de ventas en una gran empresa que fabrica una línea de interruptores, sólo ha registrado datos agrupados. Para el interruptor C9660Q, que se usa en fuentes de poder pequeñas, observa que durante el último año se embarcaron las siguientes cantidades de unidades.

Unidades embarcadas	**Número de semanas**
0-2 000	3
2 001-5 000	6
5 001-9 000	12
9 001-12 000	17
12 001-18 000	10
18 001-20 000	4

A partir de tales observaciones, estime la media y la desviación estándar de los embarques semanales. (Sugerencia: Esto se denomina datos agrupados. Para los fines de sus cálculos, suponga que todas las observaciones ocurren en el punto medio de cada intervalo.)

6. *a*) Considere el problema 3, de Oakdale Furniture. ¿Bajo qué circunstancias podría estimarse la proporción principal de consumo de pegamento?

   b) Si la demanda fuera predecible, ¿desearía usar una ley de probabilidades para describirla? ¿Bajo qué circunstancias sería viable justificar un modelo probabilístico de la demanda, aun cuando la demanda pudiera predecirse con exactitud?

5.3 EL MODELO DEL REPARTIDOR DE PERIÓDICOS

Retomemos el ejemplo 5.1. Mac quiere calcular la cantidad de ejemplares de *The Computer Journal* que debe comprar cada domingo. Al estudiar los datos históricos se vio que la demanda durante cualquier semana es una variable aleatoria que tiene una distribución normal aproximada, con media 11.73 y desviación estándar 4.74. Mac compra cada ejemplar a .25 dólares y lo vende al público a .75 dólares, y su proveedor le paga .10 dólares por cada ejemplar devuelto. Una solución obvia es que debe adquirir una cantidad suficiente para satisfacer la demanda media de aproximadamente 12 ejemplares. Hay algo equivocado en esta solución. Si Mac compra un ejemplar que no vende su gasto es $.25 - .10 = .15$, que él debe aportar. Por otro lado, si no puede satisfacer la demanda de un cliente pierde $.75 - .25 = .50$ que hubiera ganado. Por consiguiente, hay un cargo bastante mayor por no tener suficiente, en comparación con tener demasiado. Si sólo compra lo suficiente como para satisfacer la demanda media, sus existencias se agotarán con la misma frecuencia que tendrá excedentes. Nuestra intuición indica que debe comprar más que la media, pero ¿cuánto más? En esta sección se responde a ello.

Notación

Este problema es un ejemplo del modelo del repartidor de periódicos, en el que se va a pedir un solo producto al principio de un periodo, y se usa únicamente para satisfacer la demanda

durante el mismo. Suponga que todos los costos relevantes pueden calcularse con base en el inventario final. Se definen

c_o = costo del inventario positivo, por unidad, que queda al final del periodo (*costo de excedentes*).

c_u = costo de la demanda insatisfecha, por unidad. Se puede considerar como costo por unidad de inventario negativo final (*costo de faltantes*).

Supondremos que la demanda D es una variable aleatoria continua y no negativa, con función de densidad de probabilidad $f(x)$ y función de densidad acumulada $F(x)$. [En el apéndice 5-A se expone un repaso breve de la teoría de las probabilidades. En particular, ahí se definen $F(x)$ y $f(x)$.]

La *variable de decisión* Q es la cantidad de unidades que se comprarán al principio del periodo. El objetivo del análisis es calcular Q que minimiza los costos en los que se espera incurrir al final del periodo.

Deducción de la función de costo

Para analizar la mayoría de los problemas estocásticos de inventario utilizaremos el siguiente esquema general:

1. Deducir una ecuación del costo incurrido en función de la variable aleatoria D y también de la variable de decisión Q.
2. Determinar el valor esperado de esta ecuación con respecto a la función densidad, o la función de probabilidad de la demanda.
3. Calcular el valor de Q que minimice la función costo esperado.

Definiremos a $G(Q, D)$ como el costo total por exceso y faltantes incurrido al final del periodo, cuando se piden Q unidades al iniciarlo; D es la demanda. Si se compran Q unidades y D es la demanda, $Q - D$ unidades quedan al final del periodo cuando $Q \geq D$. Si $Q < D$, entonces $Q - D$ es negativo, y la cantidad de unidades que quedan disponibles al final del periodo es 0. Se observa que

$$\text{máx}\{Q - D, 0\} = \begin{cases} Q - D & \text{si } Q \geq D, \\ 0 & \text{si } Q \leq D. \end{cases}$$

De la misma manera, máx$\{D - Q, 0\}$ representa el exceso de la demanda respecto a la oferta, es decir, la demanda insatisfecha que resulta al final del periodo. Para cualquier realización de la variable aleatoria D, uno de esos términos será 0.

Por consiguiente,

$$G(Q, D) = c_o \,\text{máx}(0, Q - D) + c_u \,\text{máx}\,(0, D - Q).$$

A continuación deduciremos la función de costo esperado. Definiremos

$$G(Q) = E(G(Q, D)).$$

Si se utilizan las reglas explicadas en el apéndice 5-A para determinar el valor esperado de una función de una variable aleatoria, se obtiene

$$\begin{aligned} G(Q) &= c_o \int_0^\infty \text{máx}(0, Q - x) f(x)\,dx + c_u \int_0^\infty \text{máx}(0, x - Q) f(x)\,dx \\ &= c_o \int_0^Q (Q - x) f(x)\,dx + c_u \int_Q^\infty (x - Q) f(x)\,dx. \end{aligned}$$

Determinación de la política óptima

Para determinar el valor de Q que minimiza al costo esperado $G(Q)$ es necesario obtener una descripción adecuada de la función $G(Q)$. Tenemos que

$$\frac{dG(Q)}{dQ} = c_o\int_0^Q 1f(x)\,dx + c_u\int_Q^\infty (-1)f(x)\,dx$$
$$= c_o F(Q) - c_u(1 - F(Q)).$$

(Es un resultado de la regla de Leibniz, que indica cómo diferenciar las integrales. Esa regla se enuncia en el apéndice 5-A.)

Por consiguiente,

$$\frac{d^2G(Q)}{dQ^2} = (c_o + c_u)f(Q) \geq 0 \qquad \text{para toda } Q \geq 0.$$

Se dice que la función $G(Q)$ es *convexa* (en forma de taza) debido a que la segunda derivada es no negativa. Podemos tener una idea más clara de la forma de $G(Q)$ si la analizamos con mayor profundidad. Obsérvese que

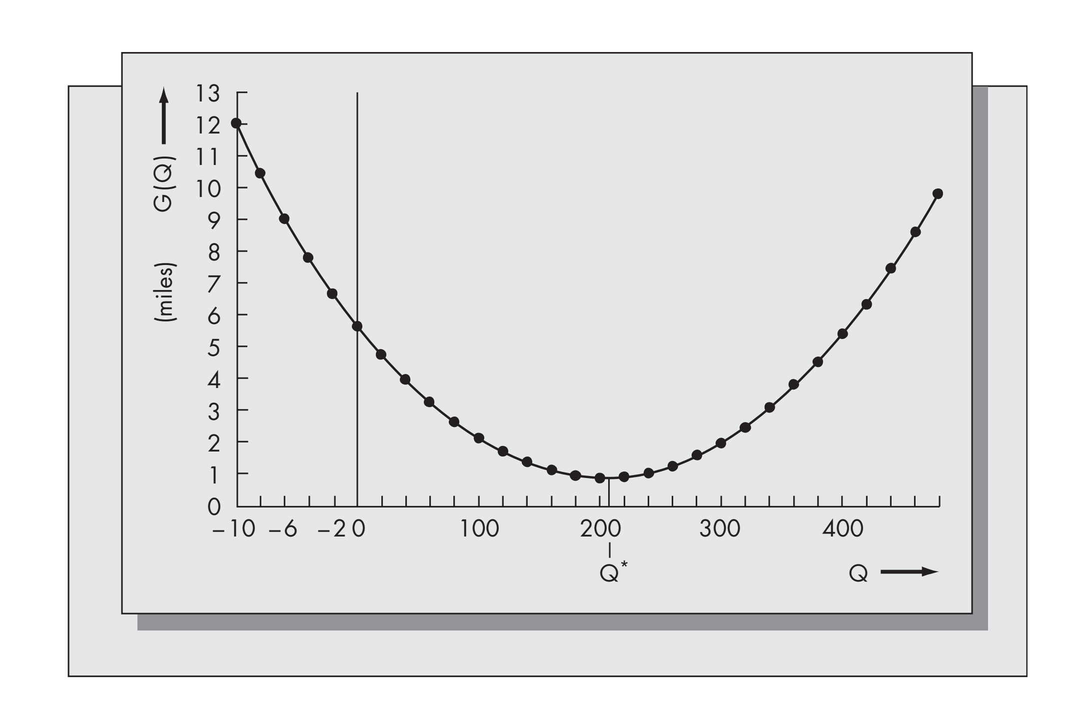
$$\left.\frac{dG(Q)}{dQ}\right|_{Q=0} = c_o F(0) - c_u(1 - F(0))$$
$$= -c_u < 0 \qquad \text{como } F(0) = 0.$$

Como la pendiente es negativa cuando $Q = 0$, $G(Q)$ es decreciente en $Q = 0$. En la figura 5-3 se grafica la función $G(Q)$.

Entonces, la solución óptima, digamos Q^*, ocurre cuando la primera derivada de $G(Q)$ es igual a cero. Esto es,

$$G'(Q^*) = (c_o + c_u)F(Q^*) - c_u = 0.$$

FIGURA 5-3
Función de costo esperado para el modelo del repartidor de periódicos.



Reordenando los términos obtenemos

$$F(Q^*) = c_u/(c_o + c_u).$$

El lado derecho de la última ecuación se denomina *relación crítica.* Como c_u y c_o son números positivos, la relación crítica está, estrictamente, entre 0 y 1. Lo anterior implica que esta ecuación tendrá una solución siempre que la demanda se distribuya en forma continua.

Dado que $F(Q^*)$ se define como la probabilidad de que la demanda no sea mayor que Q^*, la relación crítica indica cuán probable sería satisfacer toda la demanda durante el periodo si se compran Q^* unidades al inicio del mismo. Es importante comprender que esto *no* es igual a la proporción de las demandas satisfechas. Cuando los costos por faltantes y excedentes son iguales, la relación crítica equivale a un medio. En ese caso, Q^* corresponde a la *mediana* de la distribución de la demanda. La media y la mediana son iguales cuando la densidad de ésta es simétrica (como la densidad normal).

Ejemplo *5.1* *(continuación)*

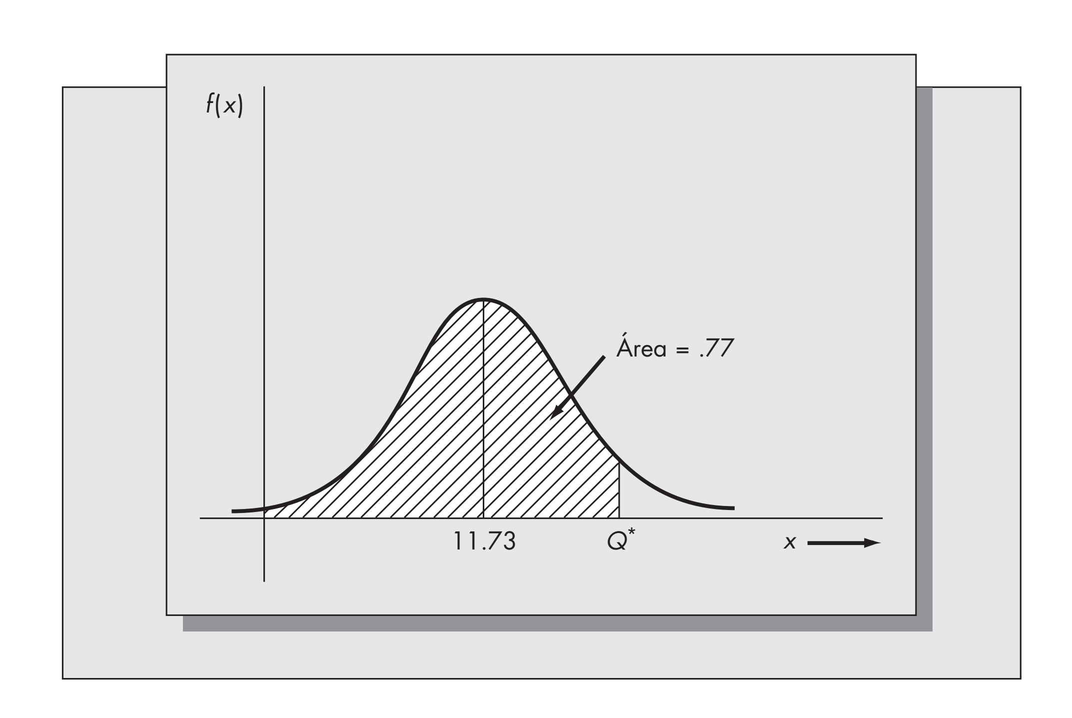
Regresemos al ejemplo del puesto de periódicos de Mac. Observamos que, según la experiencia, la demanda semanal de *Journal* tiene una distribución aproximadamente normal con media $\mu = 11.73$, y desviación estándar $\sigma = 4.74$. Como Mac compra sus revistas a .25 y puede recuperar .10 por cada ejemplar no vendido, su costo de excedente es $c_o = .25 - .10 = .15$. Su costo por faltantes es la ganancia en cada venta, y entonces $c_u = .75 - .25 = .50$. La relación crítica es $c_u/(c_o + c_u) = .50/.65 = .77$. Por consiguiente, para que la probabilidad de satisfacer toda la demanda semanal sea .77 debe comprar suficientes ejemplares. La Q^* óptima es el percentil 77 de la distribución de la demanda (véase la figura 5-4).

Usando la tabla A-1 o la tabla A-4 que se ecuentran al final del libro obtenemos un valor estandarizado de $z = .74$. La Q óptima es

$$Q^* = \sigma z + \mu = (4.74)(0.74) + 11.73$$
$$= 15.24 \approx 15.$$

Por consiguiente, debe comprar 15 ejemplares cada semana.

FIGURA 5-4
Determinación de la cantidad óptima de pedido para el ejemplo del repartidor de periódicos.



Política óptima para una demanda discreta

Nuestra deducción de la fórmula del repartidor de periódicos se basó en suponer que la demanda en el periodo está descrita por una distribución continua de probabilidades. Hicimos notar varias razones para trabajar con distribuciones continuas. Sin embargo, en algunos casos, y en especial cuando la demanda media es pequeña, quizá sería imposible obtener una representación fiel del patrón observado de la demanda si se usa una distribución continua. Por ejemplo, podría considerarse que la aproximación normal de la historia de la demanda del *Journal* durante 52 semanas, que se ilustra en la figura 5-2, no es suficientemente exacta para nuestros fines.

El procedimiento para determinar la solución óptima para el problema del repartidor de periódicos cuando se supone que la demanda es discreta es una generalización natural del caso continuo. En éste, el valor de Q en la solución óptima es el que hace que la función de distribución sea igual a la relación crítica $c_u/(c_u + c_o)$. En el caso discreto la función de distribución varía a saltos; no es probable que alguno de sus valores sea exactamente igual a la relación crítica. Esta última caerá, en el caso general, entre dos valores de $F(Q)$. El procedimiento óptimo de solución es ubicar la relación crítica entre dos valores de $F(Q)$ y elegir la Q que corresponde al valor *mayor*. (El hecho de que Q siempre se redondee hacia arriba, y que no simplemente se redondee, se puede demostrar matemáticamente con facilidad. Es distinto a suponer que las unidades están ordenadas en cantidades discretas pero la demanda es continua, en cuyo caso Q se redondea al entero más cercano, como se hizo antes.)

Ejemplo ***5.2***

Usaremos la distribución empírica establecida con un año de historia de la demanda, y no con la aproximación normal para resolver el problema del puesto de periódicos de Mac. Las probabilidades empíricas se obtienen de la figura 5-1, dividiendo entre 52 cada una de las alturas. Así se obtiene

Q	*f(Q)*	*F(Q)*	*Q*	*f(Q)*	*F(Q)*
0	1/52	1/52 (.0192)	12	4/52	30/52 (.5769)
1	0	1/52 (.0192)	13	1/52	31/52 (.5962)
2	0	1/52 (.0192)	14	5/52	36/52 (.6923)
3	0	1/52 (.0192)	15	5/52	41/52 (.7885)
4	3/52	4/52 (.0769)	16	1/52	42/52 (.8077)
5	1/52	5/52 (.0962)	17	3/52	45/52 (.8654)
6	2/52	7/52 (.1346)	18	3/52	48/52 (.9231)
7	2/52	9/52 (.1731)	19	3/52	51/52 (.9808)
8	4/52	13/52 (.2500)	20	0	51/52 (.9808)
9	6/52	19/52 (.3654)	21	0	51/52 (.9808)
10	2/52	21/52 (.4038)	22	1/52	52/52 (1.0000)
11	5/52	26/52 (.5000)			

La relación crítica para este problema fue .77, que corresponde a un valor de $F(Q)$ entre $Q = 14$ y $Q = 15$. Como se redondea hacia arriba, la solución óptima es $Q^* = 15$. Observe que es una cantidad exactamente del mismo orden que la obtenida usando la aproximación normal.

Extensión para incluir el inventario inicial

En la deducción del modelo del repartidor de periódicos supusimos que el inventario inicial en el periodo era cero. Ahora supondremos que el inventario inicial es cierto valor u, y que $u > 0$. La política óptima en este caso es una modificación sencilla del caso en que $u = 0$. Esta aplicación sería adecuada para un producto cuya vida en almacén es mayor que un periodo, aunque esta ampliación no se aplicaría a los periódicos.

Veamos la función costo esperado, $G(Q)$, ilustrada en la figura 5-3. Si $u > 0$, tan sólo significa que estamos iniciando en algún punto distinto a 0. Seguimos tratando de estar en Q^* después de ordenar, y éste sigue siendo el punto mínimo de la curva de costos. Si $u < Q^*$, esto se logra pidiendo $Q^* = u$. Si $u > Q^*$, estamos pasados de donde queremos estar en la curva. Al pedir inventario adicional tan sólo se mueve la curva de costos hacia arriba, lo cual equivale a mayores costos. Lo óptimo en este caso es simplemente no pedir.

Por consiguiente, la política óptima cuando hay un inventario inicial de $u > 0$ es

Pedir $Q^* = u$ si $u < Q^*$.

No pedir si $u \geq Q^*$.

Aplicación instantánea

UTILIZAR LOS MODELOS DE INVENTARIO PARA MANEJAR LA CADENA DE ABASTECIMIENTO DE SEMILLA DE MAÍZ EN SYNGENTA

Los agricultores plantan anualmente cientos de miles de acres de maíz en todo el mundo para cubrir la demanda de una población creciente. Compañías como Sygenta Seeds de Minnesota abastecen de semilla a dichos agricultores. Sólo en Estados Unidos, el mercado de semilla de maíz es de aproximadamente $2.3 millardos al año. Syngenta es una de las ocho compañías que representa el 73 por ciento de este mercado. ¿De dónde obtienen Sygenta y sus competidores esta semilla? La respuesta es que se produce mediante el cultivo de maíz y la cosecha de semillas. Este cultivo se realiza con el fin de cosechar la semilla y se conoce como maíz de siembra.

Resulta complicado determinar qué cantidad de maíz de siembra debe plantarse por diferentes razones. Una es que existen cientos de diferentes híbridos de semilla. Algunos funcionan mejor en climas cálidos y de mayor humedad, mientras que otros lo hacen en climas más fríos y secos. También varían el color, la textura, el contenido de azúcar, etc., del maíz producido por diferentes híbridos. Los agricultores no vuelven a utilizar ningún híbrido que no dé los resultados esperados. Por tanto, es difícil predecir la demanda anual. Además de enfrentar incertidumbre en la demanda, los productores de maíz de siembra también enfrentan rendimientos inciertos. Sus plantaciones están sujetas al mismo conjunto de riesgos que encara el resto de los agricultores: escarcha, corrientes de aire y golpes de calor.

Syngenta debe decidir cuánto maíz debe plantar cada estación. Como se desconoce la demanda de semillas, el problema suena un poco a aplicación directa del modelo del repartidor de periódicos con una demanda y un abastecimiento inciertos. Sin embargo, el problema de decisión tiene una característica adicional que lo hace más complicado que un problema normal como el del repartidor. Syngenta, al igual que muchos de sus competidores, planta maíz de siembra en los hemisferios norte y sur. Como las estaciones hemisféricas son contrarias entre sí, las plantaciones se hacen en diferentes momentos del año. En particular, el maíz se planta en primavera en cada hemisferio, así que la plantación sudamericana se realiza seis meses después de la norteamericana. Esto le da a la compañía una segunda oportunidad de incrementar los niveles de producción en Sudamérica para superar los déficit en Norteamérica, o disminuir los niveles de producción en Sudamérica cuando hay excedentes en Norteamérica.

El problema de la planificación del tamaño de plantación del maíz de siembra fue abordado por un equipo de investigadores de la Universidad de Iowa, en colaboración con un vicepresidente a cargo de abastecer a Syngenta.[1] Al utilizar aproximaciones discretas de la demanda y distribuciones de rendimiento, fueron capaces de formular el problema de planeación como un problema lineal, de manera que pudiera aplicarse a toda la compañía. Un análisis retrospectivo demostró que la compañía podía haber ahorrado hasta $5 millones usando el modelo. Los analistas también lograron identificar un sesgo sistemático en los pronósticos de semillas que, generados por la compañía, dieron como resultado una sobreproducción consistente. Actualmente utilizan el modelo matemático como una guía para sus decisiones de siembra.

[1] Jones, P. C., Kegler, G., Lowe, T. J. y Traub, R. D. "Managing the Seed-Corn Supply Chain at Syngenta", *Interfaces* 33 (1), enero-febrero de 2003, pp. 80-90.

Observe que Q^* debería ser interpretado como el punto hasta el cual debe hacerse el pedido, y no como la cantidad que se solicita cuando $u > 0$. También se llama nivel objetivo o nivel básico de existencias.

Ejemplo *5.2 (continuación)*

En el ejemplo 5.2, supondremos que Mac ha recibido 6 ejemplares del *Journal* al principio de la semana por parte de otro proveedor. La política óptima sigue siendo tener 15 ejemplares en existencia después del pedido, de modo que habría que pedir la diferencia $15 - 6 = 9$ ejemplares. (Se igualan $Q^* = 15$ y $u = 6$ para obtener la cantidad pedida $Q^* - u = 9$.)

Extensión a varios periodos de planeación

Para deducir el modelo del repartidor de periódicos se supuso que el artículo "perece" con rapidez y no puede usarse para satisfacer la demanda en los siguientes periodos. Sin embargo, en la mayoría de los ambientes industriales y de menudeo, los productos son duraderos y el inventario que queda al final de un periodo se puede almacenar y usar para satisfacer la demanda en el futuro.

Esto significa que el inventario final en cualquier periodo se transforma en el inventario inicial del periodo siguiente. En la sección anterior se indicó cómo se modifica la política óptima cuando hay inventario inicial. Sin embargo, cuando la cantidad de periodos restante es mayor que 1, también hay que modificar el valor de Q^*. En especial, la interpretación de c_o y c_u será distinta. Sólo consideraremos el caso en el que hay una cantidad infinita de periodos restantes. El valor óptimo del punto hasta el cual pedir, cuando se tiene una cantidad finita de periodos, estará entre las soluciones para un periodo y para un número infinito de periodos.

En la deducción y el análisis subsecuente de la fórmula de cantidad económica de pedido en el capítulo 4, observamos que el costo variable de pedido c sólo intervino en la optimización para determinar el costo de mantener el inventario ($h = Ic$). Además, pudo apreciarse que todas las políticas factibles de operación incurrían en el mismo costo promedio anual de reabastecimiento, λc. Sucede que lo mismo se aplica esencialmente en el problema del repartidor de periódicos con horizonte infinito. Mientras el exceso de la demanda se acumule para el siguiente periodo, en todas las políticas factibles tan sólo se pedirá la demanda durante cualquier periodo largo. De igual manera, mientras el exceso de la demanda se acumule, la cantidad de unidades vendidas tan sólo será igual a la demanda durante cualquier intervalo largo de tiempo. Por consiguiente, tanto c_u como c_o serán independientes del costo proporcional del pedido, c, y del precio de venta del artículo. En este caso se interpreta a c_u como el costo de pérdida de buena voluntad y a c_o como el costo de inventario. En el apéndice 5-B se demuestra rigurosamente que esta es la interpretación correcta de los costos de faltantes y de excedentes.

Ejemplo *5.3*

Regresemos al puesto de periódicos de Mac, descrito en los ejemplos 5.1 y 5.2. Supongamos que Mac quiere determinar cómo reabastecer el inventario de un tesauro muy popular que pide mensualmente. Los ejemplares que no vende al final del mes los conserva para futuras ventas. Supóngase que los clientes que piden ejemplares cuando se agotan esperan hasta el mes siguiente. Mac compra el tesauro a 1.15 dólares y lo vende a 2.75 dólares. Estima que el costo de pérdida de buena voluntad le cuesta .50 cada vez que hay que acumular una demanda. La demanda mensual del libro se aproxima bastante bien con una distribución normal con media de 18 y desviación estándar de 6. Mac aplica una tasa anual de interés de 20 por ciento para determinar su costo de inventario. ¿Cuántos ejemplares del tesauro debe comprar al inicio de cada mes?

Solución

El costo de excedentes en este caso es tan sólo el costo de inventario, que es $(1.15)(.20)/12 = 0.0192$ dólares. El costo de faltante es el costo de pérdida de buena voluntad, que se supone de .50. Por consiguiente, la relación crítica es $.5/(.5 + .0192) = .9630$ dólares. Según la tabla A-1 al final del libro, esto corresponde a un valor de $z = 1.79$. El valor óptimo del pedido hasta llegar al punto $Q^* = \sigma z + \mu = (6)(1.79) + 18 = 28.74 \approx 29$.

Ejemplo *5.3* *(continuación)*

Supóngase que una librería del rumbo también tiene existencias del tesauro, y que los clientes lo compran allí cuando se le terminan a Mac. En este caso, el exceso de la demanda se pierde; no se acumula. El pedido para reabastecer será distinto del obtenido cuando se supone el acumulamiento completo de la demanda. En el apéndice 5-B se demuestra que en el caso de ventas perdidas, el costo de faltantes debe interpretarse como el de la pérdida de buena voluntad más la venta perdida. El costo de excedentes debe seguirse interpretando como el costo de mantener el inventario. Por consiguiente, la solución de este ejemplo con ventas perdidas es $c_u = 0.5 + 1.6 = 2.1$ dólares. La relación crítica es $2.1/(2.1 + .0192) = .9909$ dólares, y resulta un valor de z igual a 2.36. El valor óptimo de Q en el caso de ventas perdidas es $Q^* = \sigma z + \mu = (6)(2.36) + 18 = 32.16 \approx 32$.

Si bien la solución para varios periodos parece tener la generalidad suficiente como para abarcar muchos tipos de problemas reales, padece de una limitación seria: no hay costo fijo de pedido. Esto significa que la política óptima, que es pedir hasta llegar a Q^*, requiere que el pedido se haga en cada periodo. Sin embargo, en la mayoría de los sistemas reales hay costos fijos relacionados con el pedido, y no es óptimo colocar pedidos cada periodo. Desafortunadamente, si se incluye un cargo fijo por colocar un pedido, se vuelve extremadamente difícil determinar las políticas óptimas de operación. Por esta razón se combatirá en una forma distinta el problema de la demanda aleatoria cuando hay un cargo fijo de pedido. Se supondrá que los niveles de inventario se revisan en forma continua y se desarrollará una generalización del análisis de cantidad óptima de pedido presentada en el capítulo anterior. Ese análisis se explica en la siguiente sección.

Problemas para la sección 5.3

7. Un vendedor de diarios mantiene un registro cuidadoso de la cantidad de periódicos que vende cada día, y de los diversos costos relevantes en su decisión sobre la cantidad óptima de periódicos que compra. ¿Por qué razón podrían ser erróneos sus resultados? ¿Qué necesitaría hacer para medir con exactitud la demanda diaria de periódicos?
8. Billy's Bakery produce pan francés cada mañana. La demanda diaria de su pan es una variable aleatoria estimada de acuerdo a la experiencia con la distribución siguiente

Número de panes vendidos en un día	Probabilidad
0	.05
5	.10
10	.10
15	.20
20	.25
25	.15
30	.10
35	.05

Fabricarlos le cuesta .8 dólares por pieza, y vende a 35 centavos cada uno. Los que no vende se los compra una cocina de beneficencia a 3 centavos cada uno.

a) Con base en la distribución discreta anterior, ¿cuántos panes debe hornear Billy's al principio del día? (Su respuesta debe ser un múltiplo de 5.)

b) Si tuviera usted que aproximar la distribución discreta con una distribución normal, ¿cree que la solución obtenida sea cercana a la respuesta que obtuvo en la parte *a*)? ¿Por qué?

c) Calcule la cantidad óptima de panes que debe hornear cada día con una aproximación normal. (Sugerencia: Debe calcular la media μ y la varianza σ^2 de la demanda para la distribución discreta anterior.)

9. La Crestview Printing Company imprime, una vez al año, una tarjeta de Navidad que tiene gran demanda, y la distribuye en papelerías y tiendas de regalo en todo Estados Unidos. Le cuesta .50 dólares imprimir cada tarjeta y recibe .65 dólares al venderla.

   Como las tarjetas tienen impreso el año, las que no se venden se desechan. De acuerdo con la experiencia y los pronósticos de pautas actuales de compra, la distribución estimada de probabilidades de la cantidad de tarjetas que se venderán la próxima Navidad en todo el país es

Cantidad vendida	Probabilidad
100 000-150 000	.10
150 001-200 000	.15
200 001-250 000	.25
250 001-300 000	.20
300 001-350 000	.15
350 001-400 000	.10
400 001-450 000	.05

   Calcule la cantidad de tarjetas que Crestview debe imprimir este año.

10. El distribuidor de automóviles Happy Henry's vende el EX123, un automóvil importado. Una vez cada tres meses le llega un embarque a Henry's. Puede recibir embarques de emergencia entre esos intervalos de tres meses, para reabastecerse cuando no surte la demanda con su inventario. Los embarques de emergencia tardan dos semanas, y los compradores están de acuerdo en esperar ese tiempo, pero por lo general van a otra agencia si tienen que esperar el embarque trimestral.

    De acuerdo con la experiencia, parece que la demanda del EX123 durante un intervalo de tres meses tiene distribución normal con media de 60 y varianza de 36. El costo de mantener en inventario un EX123 durante un año es 500 dólares. Los embarques de emergencia cuestan 250 dólares por automóvil, además de los costos normales de embarque.

    a) ¿Cuántos automóviles debe comprar Happy Henry's cada 3 meses?

    b) Repita los cálculos suponiendo que el exceso de demanda se acumula de un periodo trimestral al siguiente. Suponga que el costo de pérdida de buena voluntad es 100 dólares para los clientes que tienen que esperar hasta el siguiente periodo trimestral, y de 50 dólares por cliente por gastos contables.

    c) Repita los cálculos suponiendo que cuando las existencias de Happy Henry's se agotan, el cliente compra el automóvil en otra parte. En este caso, suponga que los automóviles le cuestan un promedio de 10 000 dólares a Henry's, y que los vende en un promedio de 13 500 dólares. Para estos cálculos no tome en cuenta los costos de pérdida de buena voluntad.

11. Irwin's vende determinado modelo de ventilador, y la mayor parte de las ventas las hace en los meses de verano. Hace una sola compra de ventiladores antes de cada verano, a un costo de 40 dólares por unidad, y los vende a 60 dólares cada uno. Los que no vende al final de la estación se ponen en oferta a 29 dólares, y se venden prácticamente todos. A continuación se enlistan las cantidades de ventiladores vendidos durante los últimos 10 veranos: 30, 50, 30, 60, 10, 40, 30, 30, 20 y 40.

    a) Estime la media y la varianza de la demanda de ventiladores cada verano.

    b) Suponga que la demanda de ventiladores cada verano se apega a una distribución normal, cuya media y varianza calculó usted en la parte *a*). Estime la cantidad óptima de ventiladores que debe comprar Irwin's antes de cada verano.

c) Con base en los 10 valores observados de la demanda anterior, forme una distribución empírica de probabilidades de la demanda de verano y calcule la cantidad óptima de ventiladores que debe comprar Irwin's según esa distribución.

d) De acuerdo con sus resultados en las partes *b*) y *c*), ¿diría usted que con la distribución normal se obtiene una aproximación adecuada?

12. El comprador de Needles Markup, una famosa tienda "selecta" de departamentos, debe determinar la cantidad de bolsas para dama, cuyo precio es alto, que importará de Italia para la siguiente época navideña. El costo unitario de la bolsa para el almacén es 28.50 dólares, y la venderá en 150.00 dólares. Cualquier bolsa que no venda al final de la estación se la compran unos almacenes de descuento en 20.00 dólares. Además, los contadores de Needles Markup estiman que hay un costo de .40 por cada dólar fijo en inventario, porque invertido en otra cosa habría producido una ganancia bruta. Considere que este costo sólo se aplica a las bolsas no vendidas.

a) Suponga que las ventas de las bolsas son igualmente probables para todos los valores entre 50 y 250 unidades durante la estación. Con esa base, ¿cuántas bolsas deben comprarse? (Sugerencia: Esto quiere decir que la distribución correcta de la demanda es uniforme. Este problema se puede resolver suponiendo una distribución uniforme discreta o continua.)

b) Un análisis detallado de los datos anteriores señala que la cantidad vendida de bolsas se aproxima mejor a través de una distribución normal con media de 150 y desviación estándar de 20. Ahora bien, ¿cuál es la cantidad óptima de bolsas que deben comprarse?

c) La demanda esperada fue la misma en las partes *a*) y *b*), pero las cantidades óptimas de pedido deben ser distintas. ¿A qué se debe esta diferencia?

5.4 SISTEMAS DE TAMAÑO DEL LOTE-PUNTO DE REORDEN

En el capítulo 4 se analizó la forma de la solución óptima para el modelo simple de cantidad óptima de pedido con un tiempo de demora positivo: cuando el nivel de inventario disponible llega a R, debe hacerse un pedido de Q unidades. En ese modelo la única variable independiente de decisión era Q, la cantidad de pedido. Su valor se determinó a partir de Q, λ y τ. En lo sucesivo también supondremos que la política de operación es de la forma (Q, R). Sin embargo, al generalizar el análisis CEP (cantidad óptima de pedido) para tener en cuenta la demanda aleatoria, consideraremos que Q y R son variables de decisión independientes.

El modelo del repartidor de periódicos para varios periodos no fue realista por dos razones: no incluye un costo de preparación para colocar un pedido y no tiene en cuenta un tiempo positivo de demora. Sin embargo, en la mayoría de los sistemas reales están presentes el costo de preparación y el tiempo de demora. Por esas razones, las clases de modelos que se describen en esta sección se usan con mucha mayor frecuencia en la práctica y, de hecho, forman la base de las políticas adoptadas en muchos sistemas comerciales de inventarios.

En esta sección Q es la cantidad pedida, mientras que en la sección anterior era el punto hasta el que debían llegar las existencias una vez hecho el pedido.

Haremos los siguientes supuestos:

1. El sistema tiene revisión continua. Esto es, las demandas se registran conforme suceden y el nivel del inventario disponible siempre se conoce.
2. La demanda es aleatoria y estacionaria. Eso significa que aunque no se puede predecir el valor de la demanda, el valor *esperado* durante cualquier intervalo de tiempo de longitud fija es constante. Suponga que la tasa esperada de demanda es λ unidades por año.

3. Hay un tiempo fijo de demora τ para colocar un pedido.
4. Se suponen los siguientes costos:

   Costo de preparación: K dólares por pedido.

   Costo de inventario: h dólares por unidad en inventario por año.

   Costo proporcional de pedido: c dólares por artículo.

   Costo de agotamiento: p dólares por unidad de demanda no satisfecha. A esto se le llama también costo de faltante o costo de penalización.

Descripción de la demanda

En el problema del repartidor de periódicos, la variable aleatoria adecuada es la demanda durante el periodo. Un periodo es la cantidad de tiempo necesaria para efectuar un cambio en el nivel del inventario disponible. A esto se le llama tiempo de respuesta del sistema. En el contexto de nuestro problema actual, el tiempo de respuesta es el tiempo de demora del pedido, τ. Por consiguiente, la variable aleatoria de interés es la demanda durante el tiempo de demora. Supondremos que la demanda durante el tiempo de demora es D, una variable aleatoria continua con función de densidad de probabilidad (o fdp) $f(x)$, y función de distribución acumulada (o fda) $F(x)$. Sean $\mu = E(D)$, y $\sigma = \sqrt{\text{var}(D)}$ la media y la desviación estándar de la demanda durante el tiempo de demora.

Variables de decisión

Para este problema, tenemos dos variables de decisión: Q y R, siendo Q el tamaño del lote o cantidad de pedido, y R el punto de reorden, o nivel del inventario en el momento de hacer el pedido.[3] A diferencia del modelo de cantidad económica de pedido, Q y R se consideran aquí como variables de decisión independientes. La política se implementa como sigue: cuando el nivel de inventario disponible llega a R, se hace un pedido de Q unidades que llegará dentro de τ unidades de tiempo. El funcionamiento de este sistema se representa en la figura 5-5.

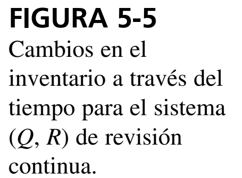
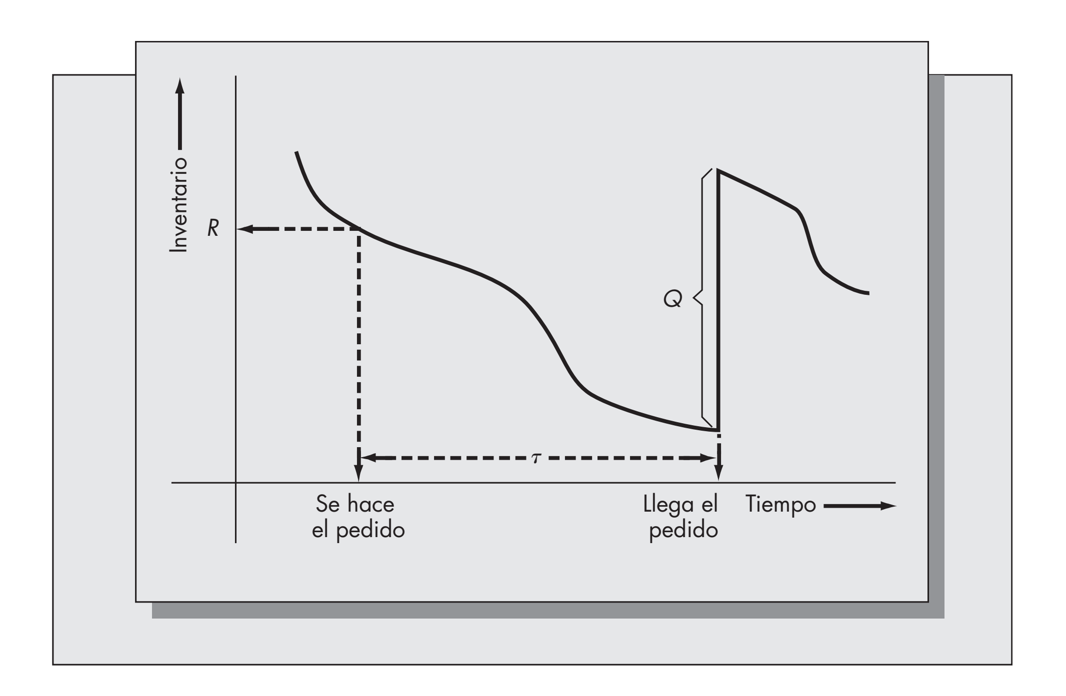


FIGURA 5-5
Cambios en el inventario a través del tiempo para el sistema (Q, R) de revisión continua.



[3] Cuando los tiempos de demora son muy largos, puede suceder que deba hacerse un pedido antes de que llegue un pedido anterior. En ese caso, la variable de decisión de reorden R debe interpretarse como la posición del inventario (el disponible más el ya pedido) cuando se coloca una *reorden*, y no como el nivel del inventario.

Deducción de la función del costo esperado

Para resolver este problema se usará el mismo método analítico que se utilizó en la deducción del modelo del repartidor de periódicos. Es decir, deduciremos una ecuación para determinar el costo anual esperado en función de las variables de decisión (Q, R) y buscaremos los valores óptimos de (Q, R) que minimicen este costo.

Costo de mantener el inventario

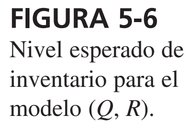
Supondremos que la tasa media de la demanda es λ unidades por año. El nivel esperado de inventario varía linealmente entre s y $Q + s$. Llamaremos a s la existencia o inventario de seguridad; se define como el nivel esperado de inventario disponible justo antes de que llegue un pedido, y se expresa con la fórmula $s = R - \lambda\tau$. La curva del nivel esperado de inventario aparece en la figura 5-6.

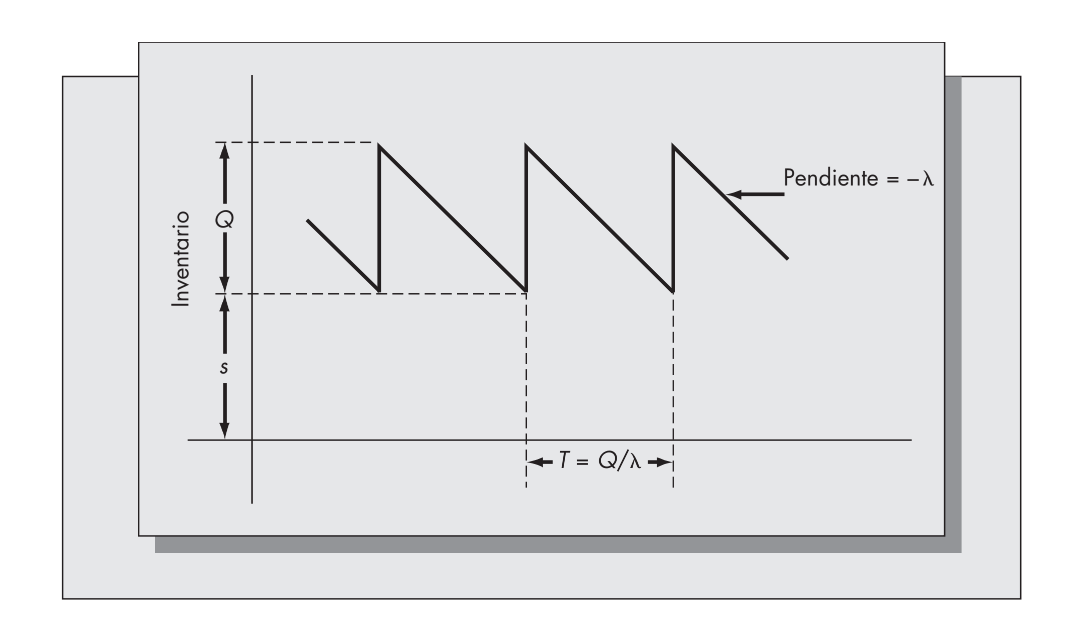
Calcularemos el costo de inventario a partir del promedio de la curva esperada. El promedio de la función que se ilustra en la figura 5-6 es $s + Q/2 = R - \lambda\tau + Q/2$. Un aspecto importante que debe hacerse notar aquí es que este cálculo sólo es una aproximación. Cuando se calcule el nivel promedio de inventario, se incluirán los casos en los que las existencias son positivas o negativas. Sin embargo, el costo de inventario *no* se debe cargar contra el inventario cuando su nivel es negativo, porque estará subestimándose el valor correcto del costo esperado de inventario. Una ecuación exacta para el promedio real de inventario es bastante complicada, y sólo se ha deducido para ciertas distribuciones específicas de la demanda. Sin embargo, en la mayoría de los sistemas reales, la proporción de tiempo con existencias agotadas es pequeño, de modo que esta aproximación suele tener una precisión razonable.

Costo de preparación

Un ciclo es el tiempo entre las llegadas de pedidos sucesivos de tamaño Q. De acuerdo con la terminología del capítulo 4, sea T la longitud esperada del ciclo. Como en el costo de preparación se incurre exactamente una vez cada ciclo, necesitamos obtener una ecuación de la longitud promedio de un ciclo para estimar con exactitud el costo de preparación por unidad de tiempo.

Existen varias formas para deducir una ecuación de la longitud esperada de ciclo. En la figura 5-6 se aprecia que la distancia entre llegadas sucesivas de pedidos es Q/λ. Otro argumento es el siguiente: la demanda esperada durante T es, naturalmente, λT. Como la cantidad de


FIGURA 5-6
Nivel esperado de inventario para el modelo (Q, R).



unidades que entran al inventario en cada ciclo es Q y hay conservación de unidades, la cantidad demandada cada ciclo debe ser, en promedio, también Q. Se obtiene el mismo resultado al igualar $Q = \lambda T$ y despejar T.

Por consiguiente, el costo promedio de preparación en el que se incurre por unidad de tiempo es $K/T = K\lambda/Q$.

Costo de penalización

En la figura 5-5 se observa que la única parte del ciclo durante la cual el sistema está expuesto a faltantes está entre el momento en que se hace un pedido y el momento en que éste llega (el tiempo de demora). La cantidad de unidades de exceso de demanda simplemente es aquella por la que la demanda durante el tiempo de demora, D, es mayor que el nivel de reorden, R. En consecuencia, la cantidad esperada de faltantes que se presentan en un ciclo está definida por la ecuación

$$E(\text{máx}(D - R, 0)) = \int_R^{\infty} (x - R)f(x)\,dx,$$

la cual se define como $n(R)$.

En esencia se trata de la misma ecuación deducida para la cantidad esperada de faltantes en el modelo del repartidor de periódicos. Como $n(R)$ representa la cantidad esperada de faltantes incurrida en cada ciclo, entonces la cantidad esperada de faltantes incurrida por unidad de tiempo es $n(R)/T = \lambda n(R)/Q$.

Componente proporcional del costo de pedido

La cantidad de unidades que entran al inventario y la que sale de él deben ser iguales durante un periodo largo. Esto significa que toda política factible debe pedir necesariamente una cantidad de unidades equivalente a la demanda sobre cualquier intervalo largo de tiempo. Esto es, toda política factible reabastecerá, en promedio, el inventario a la tasa de la demanda. Por consiguiente, el costo esperado proporcional al pedido, por unidad de tiempo, es λc. Como este término es independiente de las variables de decisión Q y R, no afecta la optimización. En adelante no lo tomaremos en cuenta.

Debe hacerse notar, sin embargo, que el costo proporcional de pedido será, en general, parte de la optimización en una forma indirecta. El costo de inventario h casi siempre se calcula multiplicando un valor adecuado de I, la tasa de interés anual, por el valor del artículo c. Por comodidad usaremos el símbolo h para representar el costo de inventario, pero téngase presente que también se podría escribir en la forma Ic.

La función de costo

Definiremos a $G(Q, R)$ como el costo anual esperado de mantener inventario, de preparación y de faltantes. Al combinar las ecuaciones deducidas para cada uno de esos términos se obtiene

$$G(Q, R) = h(Q/2 + R - \lambda\tau) + K\lambda/Q + p\lambda n(R)/Q.$$

El objetivo es elegir Q y R para minimizar $G(Q, R)$. En el apéndice 5-C presentamos los detalles de la optimización. Como se demuestra, para encontrar la solución óptima es necesario resolver de manera iterativa las dos ecuaciones siguientes:

$$Q = \sqrt{\frac{2\lambda[K + pn(R)]}{h}} \tag{1}$$

$$1 - F(R) = Qh/p\lambda. \tag{2}$$

El procedimiento de solución requiere iterar entre las ecuaciones (1) y (2) hasta que dos valores sucesivos de Q y R sean (esencialmente) iguales. El procedimiento se inicia usando

Q_0 = CEP (como se define en el capítulo 4). A continuación se calcula R_0 con la ecuación (2). Ese valor de R se usa para calcular $n(R)$, que se sustituye en la ecuación (1) para calcular Q_1, que a continuación se sustituye en (2) para obtener R_1, y así sucesivamente. La convergencia se presenta por lo general después de dos o tres iteraciones. Cuando las unidades son enteros, los cálculos deben continuarse hasta que valores sucesivos de Q y R queden a menos de una unidad de su valor anterior. Cuando las unidades son continuas, podría requerirse un criterio de convergencia de menos de una unidad, dependiendo del nivel de exactitud que se desee.

Cuando la demanda se distribuye normalmente, $n(R)$ se calcula aplicando la función estandarizada de pérdida. Esta función $L(z)$ se define de la siguiente manera

$$L(z) = \int_z^{\infty} (t - z)\phi(t)\, dt$$

en la que $\phi(t)$ es la densidad normal estándar. Si la demanda en el tiempo de demora es normal con media μ y desviación estándar σ, se puede comprobar que

$$n(R) = \sigma L\left(\frac{R - \mu}{\sigma}\right) = \sigma L(z).$$

La variable estandarizada z es igual a $(R - \mu)/\sigma$. Los cálculos de la política óptima se llevan a cabo usando la tabla A-4, al final de este libro.

Ejemplo *5.4*

Harvey's Specialty Shop es un conocido negocio que se especializa en alimentos gourmet internacionales. Uno de los artículos que vende es cierta mostaza que compra a una empresa británica. Un frasco le cuesta a Harvey's 10 dólares, y requiere un tiempo de demora de seis meses para reabastecer el inventario. Se aplica una tasa de interés anual de 20 por ciento para calcular los costos de inventario, y se estima que si un cliente pide la mostaza cuando no hay, el costo de pérdida de buena voluntad será de 25 dólares por frasco. Los gastos contables por colocar un pedido son 50 dólares. Durante el tiempo de demora de seis meses para reabasto, Harvey's calcula vender un promedio de 100 frascos, pero hay bastante variación de uno a otro periodo semestral. Se estima que la desviación estándar de la demanda durante cada periodo semestral es 25. Suponga que la demanda se describe con una distribución normal. ¿Cómo debe programar Harvey's el reabastecimiento de dicha mostaza?

Solución

Deseamos determinar los valores óptimos del punto de reorden R y del tamaño de lote Q. Para iniciar el cálculo necesitamos obtener CEP. Sin embargo, para eso se debe conocer la tasa anual de la demanda, que parece no estar especificada. Pero tengamos en cuenta que si el tiempo de demora del pedido es seis meses y la media de la demanda durante ese tiempo es 100, ello implica que la demanda anual es 200, y por tanto, $\lambda = 200$. En consecuencia, CEP $= \sqrt{2K\lambda/h} = \sqrt{(2)(50)(200)/(0.2)(10)} = 100$.

El siguiente paso es calcular R_0 con la ecuación (2). Sustituyendo $Q = 100$, obtenemos

$$1 - F(R_0) = Q_0 h/p\lambda = (100)(2)/(25)(200) = .04.$$

En la tabla A-4 vemos que el valor de z correspondiente a una cola de .04 en el lado derecho es $z = 1.75$. Se resuelve $R = \sigma z + \mu$, es decir, $R = (25)(1.75) + 100 = 144$. Además, $z = 1.75$ da como resultado $L(z) = .0162$. Por consiguiente, $n(R) = \sigma L(z) = (25)(.0162) = .405$.

Ahora podemos calcular Q_1 con la ecuación (1):

$$Q_1 = \sqrt{\frac{(2)(200)}{2}[50 + (25)(0.405)]} = 110.$$

Este valor de Q se compara con el anterior, que es 100. No se acercan lo suficiente como para detener la iteración. Sustituyendo $Q = 110$ en la ecuación (2) se obtiene $1 - F(R_1) = (110)(2)/(25)(200) = .044$. Ahora en la tabla A-4 se encuentra $z = 1.70$ y $L(z) = .0183$. Además, $R_1 = (25)(1.70) + 100 = 143$. Ahora calculamos $n(R_1) = (25)(.0183) = .4575$, y $Q_2 = = \sqrt{(200)[50 + (25)(0.4575)]} = 110.85 \approx 111$. Si

sustituimos $Q_2 = 111$ en la ecuación (2) obtenemos $1 - F(R_2) = .0444$, $z = 1.70$ y $R_2 = R_1 = 143$. Como Q_2 y R_2 están a menos de una unidad de Q_1 y R_1, podemos terminar los cálculos.

En conclusión, los valores óptimos de Q y R son $(Q, R) = (111, 143)$. Entonces, cada vez que el inventario de mostaza en Harvey's baja a 143 frascos, se debe colocar un pedido por 111 frascos.

Ejemplo *5.4* *(continuación)*

Para el mismo ejemplo, determine lo siguiente:

1. Inventario de seguridad.
2. Los costos anuales promedio de inventario, preparación y penalización asociados al control de las existencias de mostaza.
3. El tiempo promedio entre la colocación de los pedidos.
4. La proporción de ciclos de pedido en que no se presentan desabasto.
5. La proporción de las demandas que no se satisfacen.

Solución

1. El inventario de seguridad es $s = R - \mu = 143 - 100 = 43$ frascos.
2. Calcularemos por separado esos costos.

   El costo de inventario es $h[Q/2 + R - \mu] = 2[111/2 + 143 - 100] = 197$ dólares por año.

   El costo de preparación es $K\lambda/Q = (50)(200)/111 = 90.09$ dólares por año.

   El costo de faltantes es $p\lambda n(R)/Q = (25)(200)(.4575)/111 = 20.61$ dólares por año.

   En consecuencia, el costo promedio total anual asociado al control de inventario de la mostaza, suponiendo una política óptima de control, es 307.70 dólares por año.
3. $T = Q/\lambda = 111/200 = .556$ año $= 6.7$ meses.
4. En este caso necesitamos calcular la probabilidad de que no haya faltantes durante el tiempo de demora. Es igual a la probabilidad de que la demanda en el tiempo de demora no sea mayor a la del punto de reorden. Tenemos que $P\{D \leq R\} = F(R) = 1 - .044 = .956$. Llegamos a la conclusión de que no habrá agotamientos de inventario en el 95.6 por ciento de los ciclos de pedido.
5. La demanda esperada por ciclo debe ser Q (véase el argumento en la deducción del costo esperado de preparación). La cantidad que se espera de faltantes por ciclo es $n(R)$. Por consiguiente, la proporción de las demandas que no se abastecen es $n(R)/Q = .4575/111 = .004$. Otra forma de expresar este resultado es que, en promedio, se satisface el 99.6 por ciento de las demandas en el momento en que ocurren.

Para concluir esta sección, debe notarse que las ecuaciones (1) y (2) se dedujeron suponiendo que todo exceso de demanda se acumula. Esto es, cuando se demanda un artículo que no está disponible de inmediato, se satisface en un tiempo posterior. Sin embargo, en muchas situaciones de competencia, como en ventas al menudeo, un supuesto más real es que el exceso de demanda se pierde. Este caso se llama *ventas perdidas*. Mientras la probabilidad de tener un desabasto sea relativamente pequeña, las ecuaciones (1) y (2) proporcionarán soluciones adecuadas en los casos de ventas perdidas y de acumulación de pedidos. Si la probabilidad no es pequeña se necesita modificar ligeramente la ecuación (2). La versión de la ecuación (2) para ventas perdidas es

$$1 - F(R) = Qh/(Qh + p\lambda). \tag{2'}$$

El efecto de resolver las ecuaciones (1) y (2′) en forma simultánea, en lugar de (1) y luego (2), será aumentar ligeramente el valor de R y disminuir un poco el valor de Q.

5.5 NIVELES DE SERVICIO EN SISTEMAS (*Q, R*)

Con frecuencia, los gerentes pasan momentos difíciles al determinar un valor exacto del costo p de faltantes, aunque el modelo de inventario descrito es bastante realista para definir muchos sistemas. En muchos casos ese costo incluye componentes intangibles, como pérdida de buena voluntad y demoras potenciales para otras partes del sistema. Un sustituto común del costo de

faltantes es un nivel de servicio. Aunque existen varias definiciones de servicio, por lo general indica la probabilidad de que se satisfaga una demanda, o un conjunto de demandas. Los niveles de servicio se pueden aplicar a sistemas de revisión periódica y a sistemas (Q, R). En la siguiente sección se describe cómo aplicar los niveles de servicio en sistemas de revisión periódica. A continuación se explican los niveles de servicio para sistemas de revisión continua.

Enseguida se describen dos tipos de servicio, los cuales se denominan Tipo 1 y Tipo 2, respectivamente.

Servicio Tipo 1

En este caso se especifica la probabilidad de no tener faltantes durante el tiempo de demora. Usaremos el símbolo α para representar esa probabilidad. Como la especificación de α determina por completo el valor de R, se pueden desacoplar los cálculos de R y Q. El cálculo de los valores óptimos de (Q, R) sujetos a una restricción de servicio Tipo 1 es bastante directo.

a) Determine la R que satisface la ecuación $F(R) = \alpha$.

b) Iguale Q = CEP.

Se interpreta a α como la proporción de ciclos en los que no hay faltantes. Un objetivo del Tipo 1 es adecuado cuando una ocurrencia de faltantes tiene la misma consecuencia, independientemente de su tiempo o cantidad. Un ejemplo sería cuando se detiene una línea de producción al faltar 1 o 100 unidades. Por lo general, cuando decimos que nos gustaría proporcionar un servicio de 95 por ciento, estamos indicando que nos agradaría poder cumplir con el 95 por ciento de las demandas cuando se presentan, y no llenar todas las demandas en el 95 por ciento de los ciclos de pedido. Asimismo, como los diversos artículos tienen longitudes distintas de ciclo, esta medida no será consistente entre productos diferentes, y ello dificulta elegir la adecuada.

Servicio Tipo 2

Este servicio mide la proporción de las demandas que se surten con las existencias. Usaremos el símbolo β para representar esa proporción. Como se vio en la parte 5 del ejemplo 5.4, $n(R)/Q$ es la fracción promedio de las demandas que no se surten cada ciclo por agotamiento de las existencias. Por consiguiente, al especificar β se tiene como resultado la restricción $n(R)/Q = 1 - \beta$.

Esta restricción es más compleja que la que surge en el servicio Tipo 1, pues por lo general implica a Q y a R. Sucede que aunque la CEP no es óptima en este caso, comúnmente da resultados bastante buenos. Si usamos esa CEP para estimar el tamaño de lote, encontraríamos R para resolver $n(R) = \text{CEP}(1 - \beta)$.

Ejemplo *5.5*

Considere de nuevo el caso de Harvey's Specialty Shop descrito en el ejemplo 5.4. Hay incomodidad en esta empresa con el supuesto de que el costo de faltantes es 25 dólares, y decide aplicar en su lugar un criterio de nivel de servicio. Supongamos que opta por usar un objetivo de nivel de servicio igual a 98 por ciento.

1. *Servicio Tipo 1.* Si suponemos que α es .98, calculamos una R tal que $F(R) = .98$. De la tabla A-1 o la A-4 obtenemos $z = 2.05$. Al igualar $R = \sigma Z + \mu$ se obtiene $R = 151$.
2. *Servicio Tipo 2.* En este caso, $\beta = .98$. Debemos resolver la ecuación

$$n(R) = \text{CEP}(1 - \beta),$$

que equivale a

$$L(z) = \text{CEP}(1 - \beta)/\sigma.$$

Sustituyendo CEP = 100 y $\beta = .98$ se obtiene

$$L(z) = (100)(0.02)/25 = .08.$$

En la tabla A-4 a partir de las expectativas parciales unitarias normales obtenemos $z = 1.02$. Al igualar $R = \sigma z + \mu$ se obtiene $R = 126$. Observe que los mismos valores de α y β producen valores muy diferentes de R.

Para comprender con más claridad la diferencia entre esas dos medidas del servicio, considere el siguiente ejemplo. Imaginemos que hemos rastreado las demandas y los faltantes durante 10 ciclos consecutivos de pedido, y que los resultados fueron los siguientes:

Ciclo de pedido	Demanda	Faltantes
1	180	0
2	75	0
3	235	45
4	140	0
5	180	0
6	200	10
7	150	0
8	90	0
9	160	0
10	40	0

Con base en una medida de servicio Tipo 1, vemos que la fracción de los periodos en los que no hubo escasez es $8/10 = 80$ por ciento. Esto es, la probabilidad de que todas las demandas se satisfagan en un solo ciclo de pedido es .8, según estas observaciones. Sin embargo, el servicio Tipo 2 que se proporciona en dicho caso es considerablemente mejor. En este ejemplo, la cantidad total de demandas durante los 10 periodos es 1 450 (la suma de los números de la segunda columna), y la cantidad total de demandas que ocasionan faltantes es 55. Por consiguiente, la cantidad de demandas satisfechas es $1\,450 - 55 = 1\,395$. La proporción de demandas satisfechas es $1\,395/1\,450 = .9621$, aproximadamente 96 por ciento.

Se usa el término *tasa de cumplimiento* con frecuencia para describir al servicio Tipo 2, y por lo general eso es lo que los gerentes entienden por servicio. (La tasa de cumplimiento en este ejemplo fue 96 por ciento.) En este ejemplo vimos una diferencia apreciable entre la proporción de ciclos en los que todas las demandas se satisfacen (servicio Tipo 1) y la tasa de cumplimiento (servicio Tipo 2). *Aun cuando es más fácil determinar la política óptima de operación que satisface un objetivo de servicio Tipo 1, esta política no se aproximará con exactitud a un objetivo de servicio Tipo 2, y no debe usarse para sustituirlo.*

Políticas (*Q, R*) óptimas sujetas a una restricción Tipo 2

Si se usa el valor de CEP para estimar el tamaño del lote se obtienen resultados con exactitud razonable cuando se usa una restricción de tasa de cumplimiento, pero el valor CEP sólo es una aproximación al tamaño óptimo de lote. Se puede obtener un valor más exacto de Q de la siguiente manera. Veamos el par de ecuaciones (1) y (2) de donde obtuvimos los valores óptimos de Q y R cuando había costo de faltantes. Al despejar p de la ecuación (2) resulta

$$p = Qh/[(1 - F(R))\lambda],$$

que ahora puede sustituir a p en la ecuación (1). Y entonces

$$Q = \sqrt{\frac{2\lambda\{K + Qhn(R)/[(1-F(R))\lambda]\}}{h}},$$

que es una ecuación cuadrática en Q. Se puede demostrar que la raíz positiva de esta ecuación es

$$Q = \frac{n(R)}{1 - F(R)} + \sqrt{\frac{2K\lambda}{h} + \left(\frac{n(R)}{1 - F(R)}\right)^2}. \tag{3}$$

A esta ecuación la llamaremos fórmula de la cantidad de pedido para nivel de servicio, fórmula CPS.[4] Esta ecuación se resuelve en forma simultánea con

$$n(R) = (1 - \beta)Q \qquad (4)$$

para obtener valores óptimos de(Q, R) que satisfagan una restricción de servicio Tipo 2.

El lector debe notar que la versión de la ecuación (4) que se usa en los cálculos está en función de la variable estandarizada z, y está dada por

$$L(z) = (1 - \beta)Q/\sigma.$$

El procedimiento de solución es esencialmente igual al que se requiere para resolver en forma simultánea las ecuaciones (1) y (2). Se comienza con $Q_0 =$ CEP, se calcula R_0 con la ecuación (4), se usa el valor de R_0 en la ecuación (3) para calcular Q_1, y así sucesivamente, y el proceso se detiene cuando dos valores sucesivos de Q y R están suficientemente cercanos (para la mayoría de los problemas basta con que estén a menos de una unidad).

Ejemplo *5.5* *(continuación)*

Regresaremos al ejemplo 5.5. Teníamos $Q_0 = 100$ y $R_0 = 126$. Además, $n(R_0) = (.02)(100) = 2$. Usando $z = 1.02$ se obtiene $1 - F(R_0) = .154$. Continuando con los cálculos,

$$Q_1 = \frac{2}{0.154} + \sqrt{(100)^2 + \left(\frac{2}{0.154}\right)^2}$$
$$= 114.$$

Al resolver la ecuación (4) se obtiene $n(R_1) = (114)(.02)$, que equivale a

$$L(z) = (114)(0.02)/25 = 0.0912.$$

De acuerdo con la tabla A-4, $z = .95$, y así

$$1 - F(R_1) = 0.171$$

y

$$R_1 = \sigma z + \mu = 124.$$

Al avanzar en los cálculos un paso más se obtienen $Q_2 = 114$ y $R_2 = 124$. Dado que tanto Q como R están a menos de una unidad de distancia de sus valores previos, se terminan los cálculos. En consecuencia, llegamos a la conclusión de que los valores óptimos de Q y R que satisfacen una restricción de tasa de cumplimiento de 98 por ciento son $(Q, R) = (114, 124)$.

Considere el error de costo que resulta de la CEP sustituida en la fórmula CPS. Para comparar esas políticas se calculan los costos anuales promedio de mantenimiento y de preparación (obsérvese que aquí no hay costo de faltantes) para las políticas $(Q, R) = (100, 126)$ y $(Q, R) = (114, 124)$.

Recordemos las fórmulas para estos costos anuales:

$$\text{Costo de mantenimiento} = h(Q/2 + R - \mu).$$
$$\text{Costo de preparación} = K\lambda/Q.$$

Para (100, 126):

$$\left.\begin{aligned}\text{Costo de mantenimiento} &= 2(100/2 + 126 - 100) = \$152\\ \text{Costo de preparación} &= (50)(200)/100 = \$100\end{aligned}\right\}\quad \text{Total} = \$252$$

Para (114, 124):

$$\left.\begin{aligned}\text{Costo de mantenimiento} &= 2(114/2 + 124 - 100) = \$162\\ \text{Costo de preparación} &= (50)(200)/114 = \$88\end{aligned}\right\}\quad \text{Total} = \$250$$

Vemos que en este caso la aproximación CEP da como resultado costos cercanos al óptimo.

[4] La fórmula CPS también se pudo haber deducido con técnicas más convencionales de multiplicadores de Lagrange. Incluimos esta deducción para mostrar la relación entre el objetivo de la tasa de cumplimiento y el modelo del costo de faltantes.

Costo implícito de faltantes

Consideremos las soluciones de (Q, R) que obtuvimos en el ejemplo 5.5, al usar el criterio de nivel de servicio, y no el costo de faltantes. Para un servicio Tipo 2 de $\alpha = .98$ la solución fue (114, 124). Aunque no se especificó costo de faltantes, es claro que esta solución corresponde a *algún* valor de p. Esto es, hay un valor de p tal que la política (114, 124) satisface las ecuaciones (1) y (2). Este valor particular se conoce como costo implícito de faltantes.

Es fácil determinar el costo implícito de faltantes. Se despeja p en la ecuación (2) para obtener $p = Qh/[1 - F(R)\lambda]$. Este costo es una forma útil de determinar si el valor elegido de nivel de servicio es adecuado.

Ejemplo *5.5* *(continuación)*

Considere de nuevo el caso de Harvey's Specialty Shop. Con un valor de $\alpha = .98$ (servicio Tipo 1) obtuvimos la política (100, 151). El costo implícito de faltantes es $p = (100)(2)/[(.02)(200)] = 50$ dólares.

Cuando usamos un valor de $\beta = .98$ (servicio Tipo 2) obtuvimos la política (114, 124). En este caso, el costo implícito de faltantes es $p = (114)(2)/[(0.171)(200)] = 6.67$ dólares.

Escalamiento de la demanda durante el tiempo de demora

Aunque en todos los ejemplos anteriores se citó la demanda durante el tiempo de demora, en la mayoría de las aplicaciones aquélla se pronosticaría de manera periódica, por ejemplo mensualmente. En esos casos se necesitaría convertir la distribución de la demanda para que correspondiera al tiempo de demora.

Suponga que las demandas se apegan a una distribución normal. Como las sumas de variables independientes aleatorias normales también tienen distribución normal, la forma de la distribución de la demanda en el tiempo de demora es normal. Por consiguiente, lo que resta es determinar la media y la desviación estándar. Definamos a λ como la media de la demanda periódica y a ν como su desviación estándar, y sea τ el tiempo de demora en periodos. Como las medias y las *varianzas* (no las desviaciones estándar) son aditivas, la media de la demanda durante el tiempo de demora es $\mu = \lambda\tau$, y la varianza es $\nu^2\tau$. En consecuencia, la desviación estándar de la demanda durante el tiempo de demora es $\sigma = \nu\sqrt{\tau}$ (aunque la raíz cuadrada no siempre sea adecuada).[5]

Ejemplo *5.6*

La demanda semanal de cierta bujía automotriz, en un taller de reparación, tiene distribución normal con media 34 y desviación estándar 12. El tiempo de demora en su abastecimiento es seis semanas. Determine la distribución de la demanda en el tiempo de demora.

Solución

La demanda durante el tiempo de demora también se distribuye normalmente, con media $(34)(6) = 204$ y desviación estándar $(12)\sqrt{6} = 29.39$. Éstos serían los valores de μ y σ que se usarían en todos los cálculos restantes.

Estimación de sigma cuando el control de inventarios y los pronósticos están relacionados

En este capítulo, se ha supuesto hasta ahora que se conoce la distribución de la demanda durante el tiempo de demora. En la práctica se supone una forma de la distribución, pero sus parámetros deben estimarse con datos reales. Suponiendo una distribución normal para la demanda durante el tiempo de demora (que es el supuesto más común), es necesario estimar la media y

[5] Sucede con frecuencia en la práctica que hay más variación en el proceso de la demanda que el descrito por una distribución normal pura. Por esta razón, la desviación estándar de la demanda se suele expresar en la forma $\nu\tau^q$, en la que el valor correcto de q, casi siempre entre 0.5 y 1, debe determinarse para cada artículo o grupo de artículos con un análisis de datos históricos.

la desviación estándar. Cuando se dispone de la historia completa de los datos, son válidos los estimados estadísticos acostumbrados de la media y la desviación estándar, es decir, los sugeridos en la sección 5.1. Sin embargo, la mayoría de los esquemas de pronóstico *no* usan todos los datos históricos. Los promedios móviles sólo usan los últimos N valores de los datos históricos y el suavizamiento exponencial asigna factores de ponderación decrecientes a esos datos.

No queda claro en estos casos cuáles son los estimadores adecuados de la media y la desviación estándar de la demanda. Este asunto se describió en la sección 2.12. El mejor estimado de la media simplemente es el pronóstico de la demanda para el siguiente periodo. Para la varianza debe usarse el estimador para la varianza del error de pronóstico. Rara vez se comprende la lógica de lo anterior. La varianza del error de pronóstico y la de la demanda *no* son lo mismo. En el apéndice 2-A esto se estableció rigurosamente.

¿Por qué es adecuado usar la desviación estándar del error de pronóstico para estimar σ? La razón es que éste es el pronóstico que estamos usando para estimar la demanda. Las existencias de seguridad se mantienen para protegerse contra errores del pronóstico de la demanda. En general, la varianza del error de pronóstico será mayor que la de la demanda. Esto resulta del error adicional de muestreo que introduce un esquema de pronóstico que *sólo* usa una parte de los datos históricos.[6]

Al parecer el primero que recomendó usar la desviación estándar del error de pronóstico en cálculos de existencia de seguridad fue R. G. Brown (1959). El método muy usado aún hoy en día, consiste en rastrear la desviación absoluta media, DAM, del error del pronóstico con la fórmula

$$\text{DAM}_t = \alpha\, \text{DAM}_{t-1} + (1-\alpha)|F_t - D_t|$$

en donde F_t es el pronóstico de la demanda en el tiempo t y D_t es la demanda real observada en el tiempo t. El estimador de la desviación estándar del error de pronóstico en el tiempo t es 1.25 DAM_t. Si bien este método está muy difundido en los sistemas comerciales de control de inventarios, aparentemente son pocos los que advierten que al usarlo están estimando la varianza del error del pronóstico y no la varianza de la demanda, y que esas cantidades no son las mismas.

*Variabilidad del tiempo de demora

Hemos supuesto hasta ahora que el tiempo de demora τ es constante y conocido. Sin embargo, en la práctica es común la incertidumbre del tiempo de demora. Por ejemplo, el tiempo necesario para transportar materiales como petróleo, que se manejan por mar, depende de las condiciones del clima. En general, es muy difícil incorporar la variabilidad del tiempo de demora en el cálculo de las políticas óptimas de inventario. El problema es que si suponemos que los tiempos sucesivos de demora son variables aleatorias independientes, es posible que se crucen, esto es, que dos pedidos sucesivos no necesariamente se reciban en el mismo orden en el que se embarcaron.

Cuando se recurre a un solo proveedor, el cruce de pedidos no es probable. Si se quieren hacer los supuestos simultáneos de que los pedidos no se cruzan y que los tiempos sucesivos de demora son independientes, la variabilidad del tiempo de demora se puede incorporar con facilidad en el análisis. Supongamos que el tiempo de demora τ es una variable aleatoria con media μ_τ y varianza σ_τ^2. Además, supongamos que la demanda en cualquier momento t tiene

[6] En casos donde la demanda subyacente no es estacional, esto es, donde existe una tendencia o estacionalidad, la variación de la demanda debido a cambios sistemáticos puede ser más alta que la variación del error de pronóstico. No obstante, todo nuestro análisis supone patrones de demanda estacional (media constante y varianza).

media λt y varianza $\nu^2 t$. En ese caso se puede demostrar que la demanda durante el tiempo de demora tiene la media y la varianza[7]

$$\mu = \lambda\mu_\tau,$$
$$\sigma^2 = \mu_\tau\nu^2 + \lambda^2\sigma_\tau^2.$$

Ejemplo *5.7*

Harvey Gold, dueño de Harvey's Specialty Shop, pide una aceituna exclusiva de la isla Santorini, de la costa griega. A través de los años, Harvey ha observado que hay mucha variabilidad en el tiempo necesario para recibir estas aceitunas. En promedio, el tiempo de demora del pedido es cuatro meses, y la desviación estándar es seis semanas (1.5 meses). La demanda mensual de las aceitunas tiene distribución normal con media de 15 frascos y desviación estándar de 6.

Al igualar $\mu_\tau = 4$, $\sigma_\tau = 1.5$, $\lambda = 15$ y $\nu = 6$, el resultado es

$$\mu = \mu_\tau\lambda = (4)(15) = 60,$$
$$\sigma^2 = \mu_\tau\nu^2 + \lambda^2\sigma_\tau^2 = (4)(36) + (225)(2.25) = 650.25.$$

Se harían los cálculos de políticas óptimas de inventario con $\mu = 60$ y $\sigma^2 = 650.25$, como la media y la varianza de la demanda en el tiempo de demora.

Cálculos en Excel

La función estandarizada de pérdida, $L(z)$, puede calcularse en Excel. Para hacerlo, escriba $L(z)$ de la siguiente manera:

$$L(z) = \int_z^\infty (t - z)\phi(t)\,dt = \int_z^\infty t\phi(t)\,dt - z(1 - \Phi(z)) = \phi(z) - z(1 - \Phi(z)).$$

La primera igualdad resulta de la definición de la función de costo acumulado, y la segunda es una consecuencia de una propiedad bien conocida de la distribución normal estándar (véase, por ejemplo, Hadley y Within, 1963, p. 444). Para programar esta fórmula en Excel debemos utilizar la función de densidad normal estándar, que es

$$\phi(z) = \frac{1}{\sqrt{2\pi}}\exp(-0.5z^2),$$

y la función incorporada de Excel *normsdist( )*, que devuelve el valor de $\Phi(\,)$. De esta forma podemos calcular las políticas (Q, R) dentro de la hoja de cálculo. Por otro lado, podemos construir una tabla $L(z)$, que puede anidarse en una rutina de búsqueda.

Problemas para las secciones 5.4 y 5.5

13. Un almacén de refacciones automotrices tiene en inventario muchas partes que vende a las refaccionarias del vecindario. Un filtro común de aceite cuesta al almacén 1.50 dólares cada uno. Se estima que el costo de procesamiento y recibo de pedido es 100 dólares por pedido. La empresa usa un cargo de inventario basado en la tasa de interés anual de 28 por ciento.

    La demanda mensual del filtro tiene una distribución normal con media de 280 y desviación estándar de 77. Se supone que el tiempo de la demora del pedido es cinco meses.

    Suponga que si pide un filtro al almacén cuando no hay existencia del mismo, la demanda se acumula, y el costo para cada demanda acumulada se evalúa en 12.80 dólares. Calcule las siguientes cantidades:

    a) Los valores óptimos de la cantidad de pedido y el nivel de reorden.
    b) El costo anual promedio de inventario, preparación y desabasto asociado con este artículo, suponiendo que se usa la política óptima.
    c) Evalúe el costo de la incertidumbre en este proceso. Es decir, compare el costo anual promedio que obtuvo en la parte *b*) con el costo anual promedio en que hubiera incurrido si la demanda en el tiempo de la demora tuviera varianza cero.

[7] Hadley y Within (1963), p. 153.

14. Pinturas Weiss usa un sistema de inventario (Q, R) para controlar sus niveles de existencias. Para una pintura blanca, muy común, de látex, los datos históricos indican que la distribución de la demanda mensual es aproximadamente normal, con media de 28 latas y desviación estándar 8. El tiempo de demora del reabastecimiento para esta pintura es de unas 14 semanas. Cada lata le cuesta 6 dólares a Weiss. Aunque las demandas en exceso se acumulan, el dueño estima que le cuestan unos 10 dólares cada una, por gastos de contabilidad y pérdida de buena voluntad. Los costos fijos de reabastecimiento son 15 dólares por pedido, y los costos de mantener el inventario se basan en una tasa anual de interés de 30 por ciento.

    a) ¿Cuáles son los tamaños óptimos de lote y los puntos de reorden para esta pintura?

    b) ¿Cuál es la existencia óptima de seguridad para esta pintura?

15. Después de asistir a un seminario de producción, el dueño de Pinturas Weiss, el mismo del problema 14, decide que su costo de faltantes de 10 dólares puede tener errores, y cambia a un modelo de nivel de servicio. Decide establecer su tamaño de lote con la fórmula de CEP que determina su punto de reorden para que *no haya desabasto* en el 90 por ciento de los ciclos de pedido.

    a) Calcule los valores resultantes (Q, R).

    b) Suponga que, desafortunadamente, lo que quería en realidad era satisfacer el 90 por ciento de sus demandas (esto es, lograr una tasa de cumplimiento de 90 por ciento. ¿Qué tasa de cumplimiento logró realmente con la política determinada en la parte *a*)?

16. Suponga que en el problema 13 el costo de faltantes se reemplaza con un objetivo de servicio Tipo 1 igual a 95 por ciento. Calcule los valores óptimos de (Q, R) en este caso.
17. Suponga que, en el problema 13, el costo de faltantes de 12.80 dólares se sustituye con el objetivo de servicio de 95 por ciento. Calcule los valores resultantes de Q y R. Asimismo, ¿cuál es el costo implícito de faltantes en esta ocasión?
18. Suponga que el almacén usó erróneamente un objetivo de servicio Tipo 1, cuando en realidad lo que quería era uno de servicio Tipo 2 (véanse problemas 16 y 17). ¿Cuál es el costo adicional de inventario en el que se incurre cada año con este artículo a causa del error?
19. Disk Drives Limited (DDL) produce una línea de discos Winchester (discos duros) internos para microcomputadoras. Esas unidades usan placas de 3.5 pulgadas que DDL compra a un proveedor externo. Los datos de demanda y los pronósticos de venta indican que la demanda semanal de las placas parece aproximarse mucho a una distribución normal con media 38 y varianza 130. Esas placas requieren un tiempo de demora de 3 semanas para su recibo. DDL ha usado una tasa anual de interés de 40 por ciento para calcular los costos de inventario. Las placas le cuestan 18.80 dólares cada una, el costo del pedido es 75 dólares por pedido y en la actualidad la empresa usa un costo de faltantes de 400 dólares por placa. (Como este ramo es tan competido, los faltantes son muy costosos.)

    a) Debido a un convenio contractual anterior con el proveedor, DDL debe comprar las placas en lotes de 500. ¿Cuál es el punto de reorden que debe usar en este caso?

    b) Cuando DDL renegocie su contrato con el proveedor, ¿qué tamaño de lote debe incluir en el convenio?

    c) ¿Cuánta penalización, en términos de costo de preparación, inventario y desabasto, arriesga DDL al contratar pedidos por lotes demasiado grandes?

    d) El presidente de DDL no está satisfecho con el costo de faltantes de 400 dólares, y decide sustituirlo por un criterio de tasa de cumplimiento de 99 por ciento. Si DDL usara un tamaño de lote igual a la cantidad económica de pedido, ¿cuál sería su punto de reorden en este caso? También calcule el costo implícito de faltantes.

20. El restaurante Bobbi's, en Boise, Idaho, es un lugar muy concurrido para almorzar en fin de semana. Allí sirven miel de maple genuina con pan tostado y panqués. Bobbi compra la miel a una empresa de Maine, que necesita tres semanas para entregarla, y le cuesta 4 dólares la botella; puede comprarla en cualquier cantidad. Los costos fijos de pedido ascienden a unos 75 dólares por gastos de contabilidad, y los de inventario se basan en una tasa anual de 20 por ciento de interés. Bobbi estima que la pérdida de buena voluntad por parte del cliente, cuando no puede servir la miel cuando la pide, asciende a 25 dólares. Con base en su experiencia, la demanda semanal de la miel es normal, con media de 12 botellas y varianza 16. Para fines de sus cálculos, puede suponer que hay 52 semanas en un año, y que todo el exceso de demanda se acumula.

   a) ¿De qué tamaño debe ser el pedido de miel de maple y cuándo debe colocar sus pedidos Bobbi?
   b) ¿Qué nivel de servicio Tipo 1 suministra la política que determinó en la parte *a*)?
   c) ¿Qué nivel de servicio Tipo 2 suministra la política que determinó en la parte *b*)?
   d) ¿Qué política debe aplicar Bobbi si el costo de faltantes se reemplaza con un objetivo de servicio Tipo 1 igual a 95 por ciento?
   e) ¿Qué política debe aplicar Bobbi si el costo de faltantes se reemplaza con un objetivo de servicio Tipo 2 igual a 95 por ciento? (Puede usted suponer un tamaño de lote igual a la CEP.)
   f) Suponga que el proveedor de Bobbi requiere un pedido mínimo de 500 botellas. Calcule el nivel de reorden que debe usar Bobbi si desea satisfacer el 99 por ciento de las demandas de miel por parte de sus clientes.

5.6 COMENTARIOS ADICIONALES A LOS SISTEMAS DE REVISIÓN PERIÓDICA

Políticas (*s*, *S*)

En el análisis del problema del repartidor de periódicos hicimos notar que una limitación seria del modelo, desde el punto de vista práctico, es que al formularlo no se incluye un costo de preparación. El modelo (Q, R) descrito en las secciones anteriores incluía un costo de preparación de pedido, pero se suponía que los niveles de inventario se revisaban en forma continua; esto es, que se conocían siempre. ¿Cómo debe administrarse el sistema cuando hay un costo de preparación para el pedido pero los niveles de inventario sólo se conocen en puntos discretos del tiempo?

La dificultad que surge al tratar de implementar una solución con revisión continua en un ambiente de revisión periódica es que probablemente el nivel de inventario rebase el punto de reorden R durante un periodo, haciendo imposible colocar un pedido en el instante en que las existencias llegan a R. Para resolver este problema, se modifica un poco la política de operación. Se definen dos números, s y S, que se usan de la manera siguiente: cuando el nivel del inventario disponible es *menor que o igual a s*, se hace un pedido por la diferencia entre el inventario y S. Si u es el inventario inicial en cualquier periodo, entonces la política (s, S) es

Si $u \le s$, pedir $S - u$.
Si $u > s$, no pedir.

Resulta extremadamente difícil determinar los valores óptimos de (s, S), debido a que son pocos los sistemas reales de operación que los usan. Se han sugerido varias aproximaciones; una de ellas es determinar una política (Q, R) con los métodos descritos antes, e igualar $s = R$ y $S = R + Q$. En muchos casos, esta aproximación producirá resultados razonables, y quizá

sea la que se usa con más frecuencia. El lector interesado en una comparación detallada de varias políticas (s, S) aproximadas debe consultar el artículo de Porteus (1985).

*Niveles de servicio en sistemas de revisión periódica

También pueden usarse niveles de servicio cuando los niveles de inventario se revisan en forma periódica. Veamos primero un objetivo de servicio Tipo 1. Esto es, deseamos determinar el pedido hasta el punto Q para que toda la demanda se satisfaga en determinado porcentaje de los periodos. Supongamos que α es el valor del servicio Tipo 1. Entonces Q debe resolver la ecuación

$$F(Q) = \alpha.$$

Esto se debe a que $F(Q)$ es la probabilidad de que la demanda durante el periodo no sea mayor que Q. Observe que tan sólo se sustituye la relación crítica en el modelo del repartidor de periódicos por α. Para determinar la Q que satisfaga un objetivo β de servicio Tipo 2, es necesario obtener una ecuación de la fracción de las demandas que no se satisfacen cada periodo. Usando esencialmente la misma notación que en los sistemas (Q, R), se define

$$n(Q) = \int_Q^{\infty} (x - Q) f(x)\, dx.$$

Observe que $n(Q)$, que representa la cantidad esperada de las demandas que no se satisfacen al final del periodo, es igual al término que multiplica a c_u en la expresión para la función de costo esperado para el modelo del repartidor de periódicos descrito en la sección 5.3. Como la demanda por periodo es μ, entonces la proporción de las demandas que no se satisfacen cada periodo es $n(Q)/\mu$. Así, el valor de Q que cumple con un objetivo de tasa de cumplimiento de β resuelve a

$$n(Q) = (1 - \beta)\mu.$$

La especificación de un objetivo de servicio Tipo 1, o uno Tipo 2, determina por completo la cantidad de pedido, independientemente de los parámetros de costo.

Ejemplo *5.8*

Mac, el dueño del puesto de periódicos del ejemplo 5.1, desea emplear un nivel de servicio Tipo 1 igual a 90 por ciento para controlar su reabasto de *The Computer Journal*. El valor de z que corresponde al 90 por ciento de la distribución normal es $z = 1.28$. Por consiguiente,

$$Q^* = \sigma z + \mu = (4.74)(1.28) + 11.73 = 17.8 \approx 18.$$

Con un servicio Tipo 2 de 90 por ciento obtenemos

$$n(Q) = (1 - \beta)\mu = (0.1)(11.73) = 1.173.$$

Por consiguiente, $L(z) = n(Q)/\sigma = 1.173/4.74 = .2475$. De acuerdo con la tabla A-4, al final del libro, se ve que

$$z \approx 0.35;$$

entonces

$$Q^* = \sigma z + \mu = (4.74)(0.35) + 11.73 = 13.4 \approx 13.$$

Observe que, al igual que en los modelos (Q, R), hay una diferencia notable entre los valores de Q^* que resultan para los mismos niveles de los servicios Tipo 1 y Tipo 2.

Problemas para la sección 5.6

21. Considere el caso de la Crestview Printing Company descrito en el problema 9. Suponga que la empresa desea producir tarjetas suficientes como para abastecer toda la demanda

navideña con 90 por ciento de probabilidad. ¿Cuántas tarjetas debe imprimir? Suponga que la probabilidad es 97 por ciento. Entonces, ¿cuántas debe imprimir? (Su respuesta dependerá del supuesto que haga acerca de la forma de la función de distribución acumulada.)

22. Regrese al caso del vendedor de automóviles Happy Henry's que se describió en el problema 10.

   a) ¿Cuántos EX123 debe comprar Happy Henry's para satisfacer toda la demanda durante un intervalo de tres meses, con 95 por ciento de probabilidad?

   b) ¿Cuántos automóviles debe comprar si el objetivo es satisfacer el 95 por ciento de las demandas?

23. Para el problema de control del inventario de pintura blanca de látex en la tienda de Weiss, descrito en el problema 14, suponga que la pintura se pide en base mensual, y no en base continua.

   a) Usando la solución (Q, R) que obtuvo en la parte *a*) del problema 14, calcule los valores adecuados de (s, S).

   b) Suponga que las demandas durante los meses de enero a junio fueron

Mes	Demanda	Mes	Demanda
Enero	37	Abril	31
Febrero	33	Mayo	14
Marzo	26	Junio	40

   Si el inventario inicial de enero fue 26 latas de pintura, calcule la cantidad de latas que se piden cada mes, de enero a junio, adoptando la política (s, S) que determinó en la parte *a*).

5.7 SISTEMAS MULTIPRODUCTO

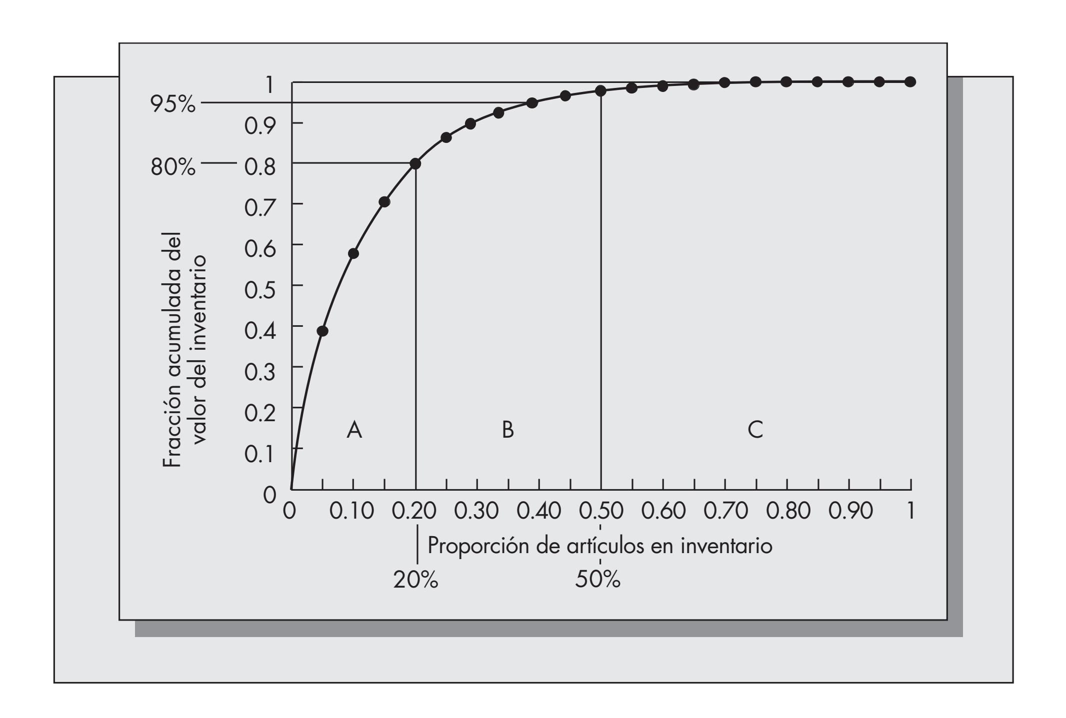
Análisis ABC

Un asunto que no hemos comentado es el costo de implementar un sistema de control de inventarios, y los intercambios entre el costo de controlar el sistema y los beneficios potenciales obtenidos con ese control. En los sistemas de inventario con varios productos, no todos tienen igual rentabilidad. Los costos de control se pueden justificar en unos casos y en otros no. Por ejemplo, es claro que no resulta económico gastar anualmente 200 dólares para monitorear un artículo que sólo aporta 100 dólares de ganancia anual.

Por esta razón es importante diferenciar entre los artículos rentables y los no rentables. Para hacerlo tomaremos prestado un concepto de la economía. Vilfredo Pareto fue un economista que estudió la distribución de la riqueza en el siglo XIX, y notó que una gran parte de ésta pertenece a un segmento pequeño de la población. Este *efecto Pareto* también se aplica a los sistemas de inventario: una gran parte del volumen monetario total de las ventas se debe, con frecuencia, a una pequeña cantidad de artículos en el inventario. Supongamos que los artículos se clasifican en orden decreciente por sus valores monetarios de ventas anuales. El valor *acumulado* de las ventas da como resultado, en general, una curva muy parecida a la de la figura 5-7.

En forma típica, el 20 por ciento superior de los artículos explica el 80 por ciento del volumen monetario anual de las ventas; el siguiente 30 por ciento, el 15 por ciento de las ventas, y el 50 por ciento restante, el último 5 por ciento del valor monetario. Esas cifras sólo son

FIGURA 5-7
Curva de Pareto: distribución de inventario por valor.



aproximadas, y varían ligeramente de un sistema a otro. Los tres grupos de artículos se identifican como A, B y C, respectivamente. Cuando se necesita una distinción más refinada se pueden usar cuatro o cinco categorías. Aun cuando sólo se usen tres categorías, los porcentajes usados para definir los artículos A, B y C podrían ser distintos a los porcentajes de 80, 15 y 5 que acabamos de recomendar.

Como los artículos A tienen la tajada más grande en las ganancias anuales, deben ser vigilados más estrechamente. Los niveles de inventario para los artículos A deben monitorearse en forma continua. Se podrían usar procedimientos más complejos de pronóstico, y debe tenerse más cuidado al estimar los diversos parámetros de costo necesarios para establecer las políticas de operación. Para los artículos B, los inventarios podrían revisarse en forma periódica, se podrían pedir grupos de artículos y no artículos individuales, y podrían usarse métodos de pronóstico menos complicados. El grado mínimo de control se aplicaría a los artículos C. Para artículos C muy poco costosos con niveles moderados de demanda, se recomiendan tamaños grandes de lote para minimizar la frecuencia con que se piden. Para los artículos C costosos con muy poca demanda, la mejor política es no tener existencias. Tan sólo se piden esos artículos cuando se demandan.

Ejemplo *5.9*

Se selecciona al azar una muestra de 20 artículos distintos de la bodega de Harvey's Specialty Shop. Los precios de esos artículos varían de .25 a 24.99 dólares, y su demanda promedio anual varía de 12 a 786. Los resultados del muestreo se registran en la tabla 5-1. En la tabla 5-2 se categorizan en orden decreciente por el valor monetario de sus ventas. Observe que sólo 4 de los 20 artículos almacenados explican más del 80 por ciento del volumen anual monetario generado por todo el grupo. También observe que hay artículos costosos en las categorías A y C.

Este informe fue muy útil para Harvey, quien había supuesto que un queso de cabra empacado del sur de Francia, el R077, era rentable debido a su costo, y no se había percatado del TTR77, un chocolate en barra del país.

TABLA 5-1 **Comportamiento de 20 artículos de almacén seleccionados al azar**

Número de parte	Precio en dólares	Demanda anual	Volumen en dólares
5497J	$ 2.25	260	$ 585.00
3K62	2.85	43	122.55
88450	1.50	21	31.50
P001	0.77	388	298.76
2M993	4.45	612	2 723.40
4040	6.10	220	1 342.00
W76	3.10	110	341.00
JJ335	1.32	786	1 037.52
R077	12.80	14	179.20
70779	24.99	334	8 346.66
4J65E	7.75	24	186.00
334Y	0.68	77	52.36
8ST4	0.25	56	14.00
16113	3.89	89	346.21
45000	7.70	675	5 197.50
7878	6.22	66	410.52
6193L	0.85	148	125.80
TTR77	0.77	690	531.30
39SS5	1.23	52	63.96
93939	4.05	12	48.60

TABLA 5-2 **Veinte artículos de almacén clasificados en orden decreciente del volumen monetario anual**

Número de parte	Precio en dólares	Demanda anual	Volumen en dólares	Volumen acumulado en dólares	
70779	$24.99	334	$8 346.66	$ 8 346.66	Artículos A: 20% de los artículos suman 80.1% del valor total.
45000	7.70	675	5 197.50	13 544.16	
2M993	4.45	612	2 723.40	16 267.56	
4040	6.10	220	1 342.00	17 609.56	
JJ335	1.32	786	1 037.52	18 647.08	Artículos B: 30% de los artículos suman 14.8% del valor total.
5497J	2.25	260	585.00	19 232.08	
TTR77	0.77	690	531.30	19 763.38	
7878	6.22	66	410.52	20 173.90	
16113	3.89	89	346.21	20 520.11	
W76	3.10	110	341.00	20 861.11	
P001	0.77	388	298.76	21 159.87	Artículos C: 50% de los artículos suman 5.1% del valor total.
4J65E	7.75	24	186.00	21 345.87	
R077	12.80	14	179.20	21 525.07	
6193L	0.85	148	125.80	21 650.87	
3K62	2.85	43	122.55	21 773.42	
39SS5	1.23	52	63.96	21 837.38	
334Y	0.68	77	52.36	21 889.74	
93939	4.05	12	48.60	21 938.34	
88450	1.50	21	31.50	21 969.84	
8ST4	0.25	56	14.00	21 983.84	

Curvas de intercambio

Gran parte de nuestro análisis supone que hay un solo artículo aislado, y que los parámetros relevantes de costo, K, h y p (o sólo K y h en el caso de niveles de servicio) son constantes con valores "correctos" que pueden determinarse. Sin embargo, convendría imaginar que uno o todos los parámetros de costo son variables de política. Los valores correctos son aquellos que

dan como resultado un sistema de control con características que satisfacen las necesidades de la empresa y los objetivos de la administración. En un sistema normal de varios productos se usan los mismos valores de costo K de preparación y de tasa de interés I. Podemos considerar que la relación K/I es una variable de política; si esa relación es grande, los tamaños de lote serán mayores y la inversión promedio en inventarios será mayor. Si esa relación es pequeña, aumentará la cantidad de reabastecimientos anuales.

Para apreciar con exactitud cómo se traza una curva de intercambio, veamos un sistema determinista formado por n productos cuyas tasas de demanda son variables, iguales a $\lambda_1, \ldots,$ λ_n, valores de artículos semejantes a $c_1, \ldots, c_n$. Si se usan los valores de la cantidad económica de pedido (CEP) para reabastecer las existencias de cada artículo, entonces

$$Q_i = \sqrt{\frac{2K\lambda_i}{Ic_i}} \qquad \text{para } 1 \leq i \leq n.$$

Para el artículo i, el tiempo del ciclo es Q_i/λ_i, de modo que λ_i/Q_i es la cantidad de reabastecimientos en un año. La cantidad total de reabastecimientos del sistema total es $\Sigma\lambda_i/Q_i$. El inventario promedio disponible del artículo i es $Q_i/2$ y el valor monetario de ese inventario es $c_iQ_i/2$. Por consiguiente, el valor total del inventario es $\Sigma c_iQ_i/2$.

Cada elección de la relación K/I dará como resultado un valor distinto de cantidad de reabastecimientos por año, y de valor monetario del inventario. Al variar K/I se puede trazar una curva como la que se ilustra en la figura 5-8. Una curva de intercambio como esta permite que la administración vislumbre con facilidad el intercambio entre el valor de la inversión en inventario y la frecuencia del reabastecimiento de existencias.

También se pueden usar las curvas de intercambio para comparar diversas estrategias de inventario de seguridad y de nivel de servicio. Por ejemplo, imaginemos un sistema en el que se aplica una restricción de tasa de cumplimiento (es decir, un servicio Tipo 2) para todos los artículos. Además, supongamos que la distribución de la demanda en el tiempo de demora para todos los artículos es normal, y que a cada artículo se le da un mismo servicio. El valor monetario del inventario de seguridad es $\Sigma c_i(R_i - \mu_i)$, y el valor anual de la demanda acumulada es $\Sigma c_i\lambda_i n(R_i)/Q_i$. Un valor fijo de β, la tasa de cumplimiento, dará como resultado un conjunto de valores de las variables de control $(Q_1, R_1), \ldots, (Q_n, R_n)$, que se puede calcular con los métodos descritos antes en este capítulo. Cada conjunto de valores de (Q, R) produce un par de

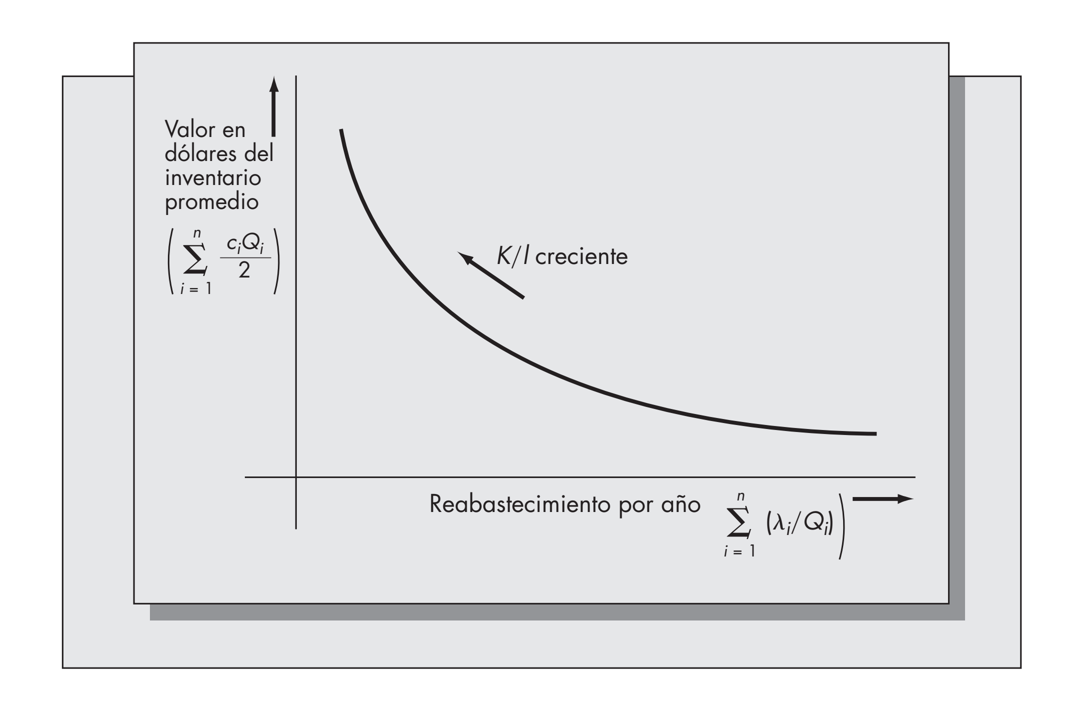
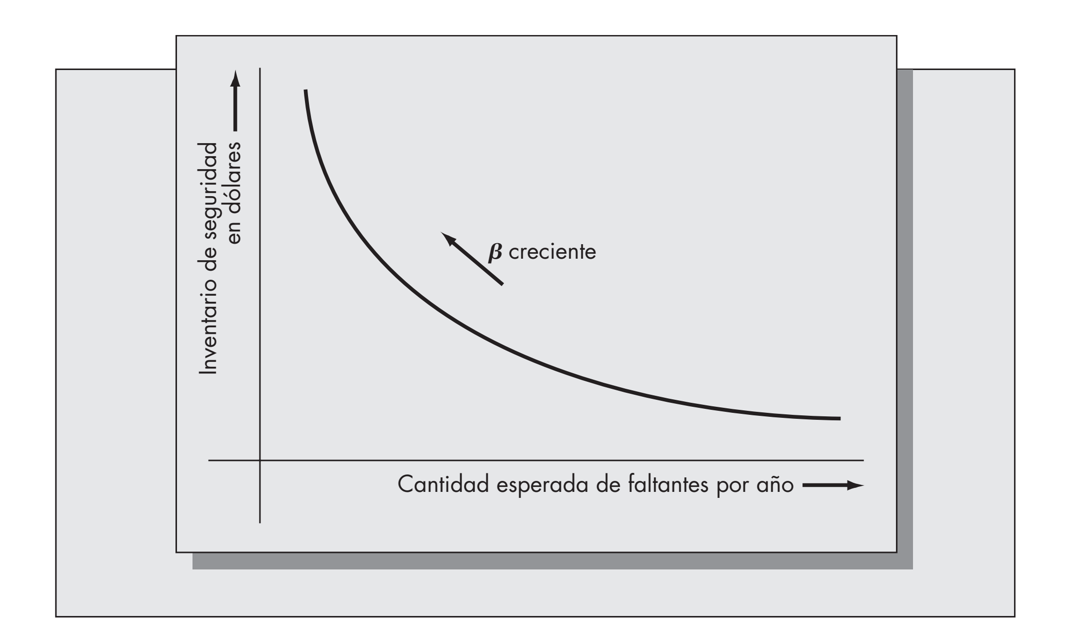


FIGURA 5-8
Curva de intercambio para la frecuencia de reabastecimiento y el valor de los inventarios.

FIGURA 5-9
Curva de intercambio para la inversión en existencias de seguridad y β.

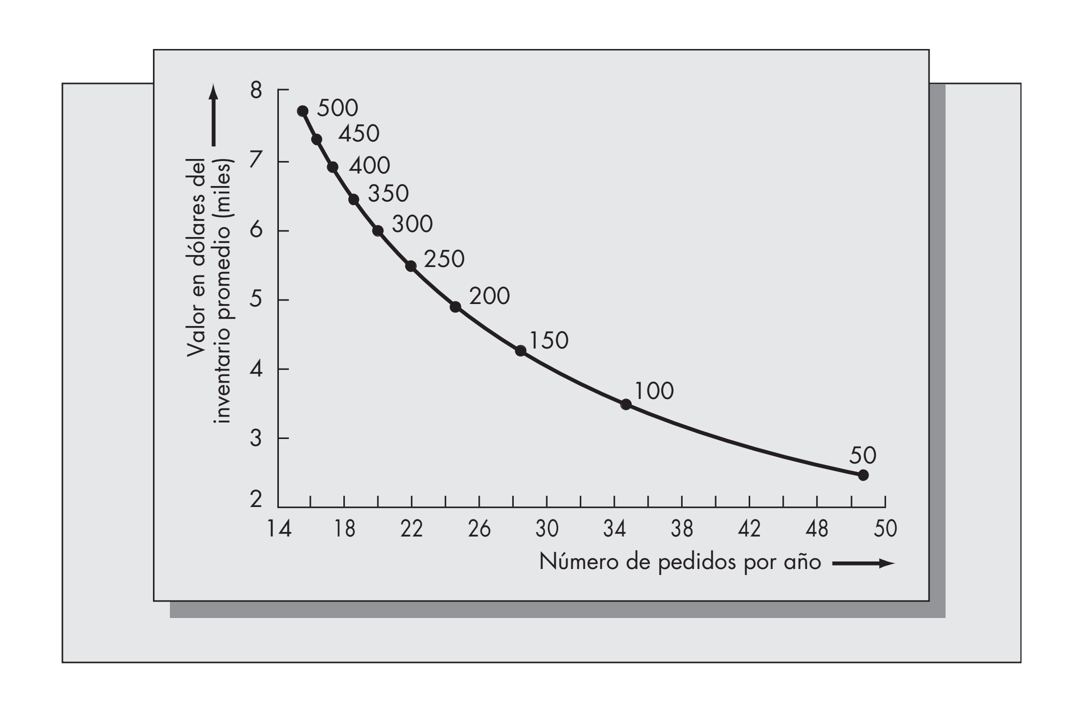


valores de existencia de seguridad y de demanda acumulada. A medida que aumenta la tasa de cumplimiento, se incrementa la inversión en inventario de seguridad y disminuye el valor de la demanda acumulada. La curva de intercambio que se podría obtener se ilustra en la figura 5-9. Tales curvas de intercambio son útiles para que la administración evalúe el impacto monetario de diversos niveles de servicio.

Ejemplo *5-10*

Regresemos a los 20 artículos de almacén referidos en las tablas 5-1 y 5-2. Supongamos que Harvey, dueño de Harvey's Specialty Shop, reconsidera sus definiciones de costos de preparación de 50 dólares y de la tasa de interés de 20 por ciento. Él usa la fórmula de CEP para calcular tamaños de lote de los 20 artículos, para un intervalo de valores de K/I de 50 a 500. La curva de intercambio que obtiene aparece en la figura 5-10.

FIGURA 5-10
Curva de intercambio: Harvey's Specialty Shop.



Harvey está operando en $K/I = 50/.2 = 250$, que da como resultado aproximadamente 22 pedidos por año y un costo promedio anual de inventario de 5447 dólares. Si reduce K/I a 100, el costo de inventario para esos 20 artículos se reduce a 3 445 dólares, y su frecuencia aumenta a 34 pedidos por año. Después de pensarlo, Harvey decide que los gastos adicionales en tiempo y contabilidad que se requieren para rastrear 12 pedidos adicionales al año justifican, definitivamente, los 2000 dólares de ahorro en costos de inventario. (Le parece muy razonable la tasa de 20 por ciento de interés, lo que significa que el valor real de su costo de preparación para los pedidos se acerca más a 20 dólares que a 50. De esta forma, la curva de intercambio puede ayudar en la determinación del valor correcto de los parámetros de costo, que de otra forma sería difícil estimar.) También considera la alternativa de pasar a $K/I = 50$, pero decide que los ahorros adicionales, de unos 1000 dólares, no valen la pena en comparación a tener que procesar casi 50 pedidos por año.

Problemas para la sección 5.7

24. Describa el sistema de clasificación ABC. ¿Cuál es el propósito de clasificar los artículos en esta forma? ¿Cuál sería el valor principal del análisis ABC para un minorista? ¿Y para un fabricante?
25. Una tienda de regalos vende al menudeo los artículos de la siguiente lista.

Artículo	Volumen anual	Ganancia promedio por artículo
Tarjetas de felicitación	3 870	$0.40
Camisetas	1 550	1.25
Joyería para caballeros	875	4.50
Novedades	2 050	12.25
Ropa para niños	575	6.85
Galletas de chocolate	7 000	0.10
Aretes	1 285	3.50
Otros artículos de joyería	1 900	15.00

a) Clasifique las categorías de artículos en orden decreciente de ganancias anuales. Clasifique cada uno en las categorías A, B o C.

b) ¿Por qué razón optaría el propietario de la tienda por vender las galletas de chocolate aun cuando podrían ser su renglón menos rentable?

26. Desde la perspectiva de la administración, ¿cuál es el valor principal de una curva de intercambio? Describa la curva de intercambio para la frecuencia de reabastecimiento y de valor de inventario, y la curva de intercambio para la cantidad esperada de faltantes por año y la inversión en existencias de seguridad.
27. Considere los ocho artículos de la lista del problema 25 que se presentan a continuación. Suponga que los costos promedio de esas categorías son

Artículo	Costo
Tarjetas de felicitación	$ 0.50
Camisetas	3.00
Joyería para caballeros	8.00
Novedades	12.50
Ropa para niños	8.80
Galletas de chocolate	0.40
Aretes	4.80
Otros artículos de joyería	12.00

Compare la cantidad total de reabastecimientos por año y el valor en efectivo del inventario de esos artículos para los siguientes valores de la relación K/I: 100, 200, 500, 1 000. A partir de los cuatro puntos obtenidos, estime la curva de intercambio de frecuencia de abastecimiento y de valor de inventario.

*5.8 PERSPECTIVA DE TEMAS AVANZADOS

Este capítulo sólo describió una pequeña parte de los modelos de inventario disponibles. Se ha realizado mucha investigación en torno al análisis de problemas mucho más complejos de control de inventarios estocástico, pero la mayoría de esas investigaciones están fuera del ámbito de nuestra exposición. Esta sección presenta una breve perspectiva de dos áreas que no se citaron en este capítulo, y que son objeto de estudio de una gran parte de los recientes trabajos sobre administración de inventarios.

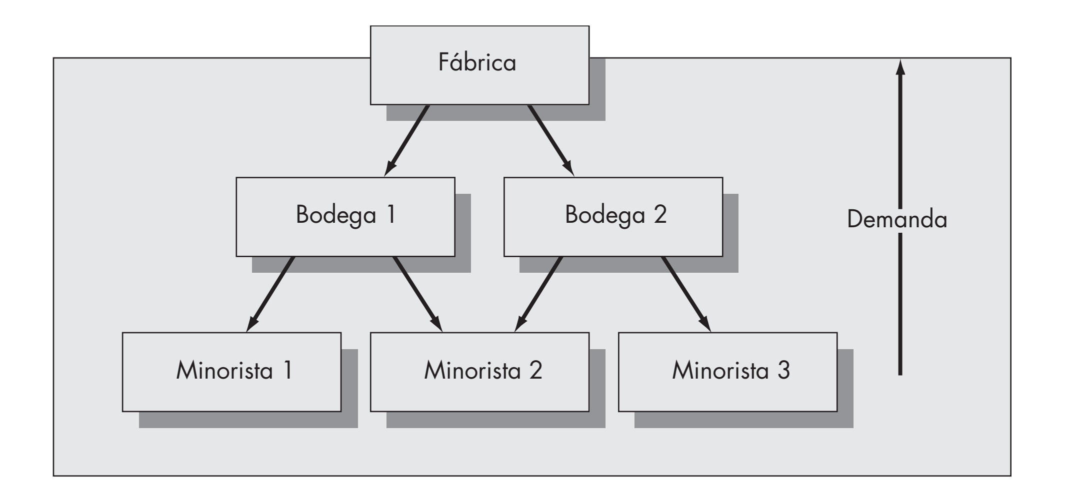
Sistemas multiniveles

Las empresas que se dedican a la manufactura y distribución de productos al consumidor deben tener en cuenta las interacciones de los diversos niveles en la cadena de distribución. En forma típica, los artículos se producen en una o más fábricas, se envían a bodegas de almacenamiento intermedio y después se embarcan a las tiendas al menudeo, donde se venden al consumidor. A los escalones del sistema los llamaremos *niveles*. Por ejemplo, el sistema de tres niveles fábrica-bodega-consumidor se representa en la figura 5-11.

Un sistema como el ilustrado en la figura 5-11 se llama, por lo general, sistema de *distribución*. En él, la demanda surge en el último nivel (en este caso el detallista) y se transmite hacia los niveles superiores. Los planes de producción en la fábrica deben coordinarse con los pedidos colocados por los detallistas en las bodegas. Otro tipo de sistema multiniveles que se maneja en contextos de manufactura se llama sistema de *ensamble*. En él, los componentes se combinan para formar subgrupos, que al unirse crean artículos finales (esto es, productos finales). Un sistema típico de ensamble se representa en la figura 5-12.

En dicho sistema, la demanda externa se origina sólo en el nivel de artículo final. La demanda de los componentes sólo se produce cuando se colocan “pedidos” a mayores niveles en el sistema, como consecuencia de los programas de producción de artículos finales. La metodología de planeación de requerimientos de materiales se diseña para modelar sistemas de este tipo, pero no tiene en cuenta el efecto de la incertidumbre sobre la demanda final. Aunque se ha realizado investigación sobre este problema, aún está por definirse la manera exacta en que la

FIGURA 5-11
Sistema típico de distribución con tres niveles.



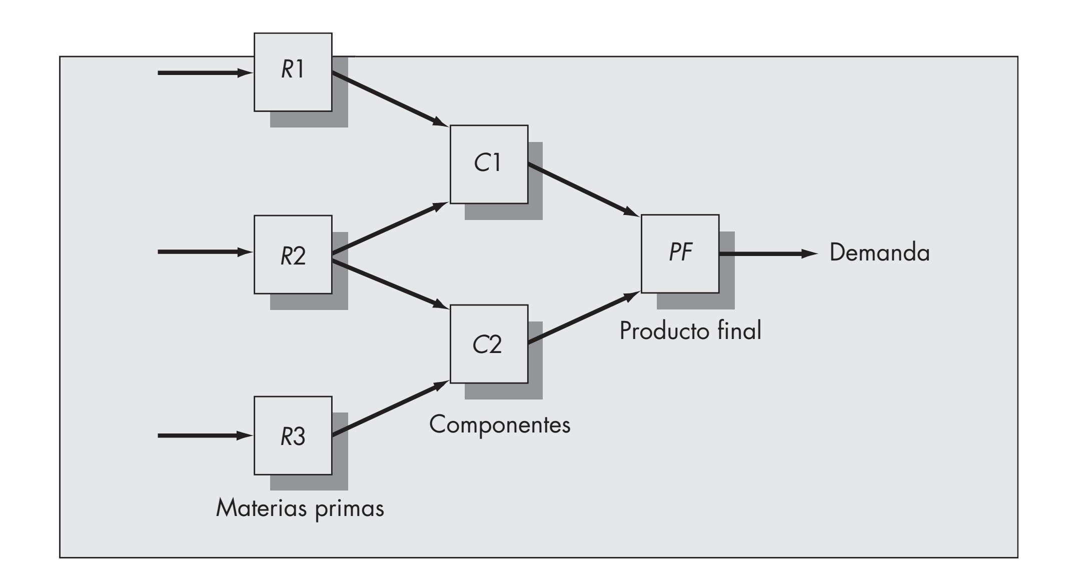


FIGURA 5-12
Sistema típico de ensamble con tres niveles.

incertidumbre en la demanda final afecta a las políticas óptimas de producción y reabasto de los componentes.

Uno de los sistemas multinivel que ha recibido gran atención en las publicaciones (Sherbrooke, 1968) fue diseñado para aplicarse en la logística militar. Imaginemos el problema de reabastecimiento que se presenta cuando se mantiene un suministro de refacciones para respaldar las reparaciones de motores de reacción para aviones. Cuando un motor (o uno de sus componentes) falla, se envía a un taller central para que lo reparen. El taller mantiene su propio inventario de motores, y puede embarcar una refacción a la base antes de terminar la reparación. El problema es determinar los niveles correctos de repuestos en la base y en el taller, teniendo en cuenta los objetivos opuestos de limitaciones de presupuesto y la posibilidad de tener una alta tasa de cumplimiento.

La inversión estadounidense en partes de repuesto es enorme. Ya en 1968 se estimaba que sólo la Fuerza Aérea tenía invertidos 10 000 millones de dólares en inventarios de artículos reparables. Teniendo en cuenta los efectos de la inflación y los tamaños de los demás servicios armados, es probable que la inversión actual en artículos reparables en las fuerzas estadounidenses sea mayor a 100 000 millones de dólares. Los modelos matemáticos de inventario han tenido gran impacto en la administración de estos inventarios, tanto militares como no militares.

En el sector privado, el menudeo es un área en la que son comunes los sistemas multinivel de distribución. La mayoría de las cadenas de ventas al menudeo usan un centro de distribución como punto de almacenamiento intermedio, entre el fabricante y la tienda minorista. Determinar cómo deben asignarse inventarios entre el centro de distribución y las tiendas es un asunto estratégico importante. El inventario en los centros permite “agrupamiento de riesgo” entre las tiendas y facilita la redistribución de inventarios de las tiendas que pudieran inclinarse hacia el desequilibrio. Hay varios paquetes de inventarios para computadora disponibles en torno a la administración de inventarios al menudeo; quizá el más conocido sea Inforem, diseñado y vendido por IBM. Los lectores interesados en una reseña detallada del problema de administración de inventarios para menudeo pueden consultar el artículo de Nahmias y Smith (1992).

Problemas con inventarios perecederos

Aunque se pueden considerar ciertos tipos de deterioro en los ajustes del costo de mantenimiento de un inventario, los modelos matemáticos que describen lo perecedero, el deterioro o la obsolescencia son, por lo general, muy complicados.

Entre los ejemplos típicos de inventarios de artículos perecederos con vida fija se incluyen los alimentos, la sangre, las medicinas y la película fotográfica. Para rastrear la edad de las unidades almacenadas es necesario conocer la cantidad de cada nivel de edad en el inventario disponible. Esto da como resultado un complicado espacio multidimensional de estados del sistema. La calidad de los artículos perecederos agrega una penalización adicional en el mantenimiento del inventario y reduce así el tamaño de los pedidos. La aplicación más notable de la teoría de los modelos de inventario perecedero ha sido en la administración de los bancos de sangre.

Un problema algo más sencillo, pero relacionado, es el del deterioro exponencial, en el que se supone que en cada periodo se pierde una fracción fija del inventario. El deterioro exponencial es una descripción fiel de los cambios que ocurren en líquidos volátiles, como alcohol y gasolina, y es un modelo preciso para la declinación de la radiactividad de los radiofármacos y de los combustibles nucleares.

Se ha modelado la obsolescencia suponiendo que la longitud del horizonte de planeación es una variable aleatoria con distribución conocida de probabilidad. En la práctica esos modelos son valiosos sólo si la distribución de la vida útil del artículo se puede estimar de manera adecuada.

Un problema relacionado es la administración de bienes estilísticos, como las modas en la industria del vestido. El problema de los bienes estilísticos difiere de los modelos de inventario descritos en este capítulo en que no sólo es incierta la demanda misma, sino también su distribución. Esto es característico de los problemas para los que no hay historia de demanda. Los procedimientos sugeridos para el problema de los bienes estilísticos implica, por lo general, algún esquema bayesiano de actualización, para combinar las observaciones actuales de la demanda con estimados anteriores de la distribución de la demanda.

5.9 INTERCAMBIO ELECTRÓNICO DE DATOS

El *intercambio electrónico de datos* (IED) constituye un avance importante en el control de inventarios y en la administración de la cadena de suministro, que cada vez gana mayor aceptación. IED es el término genérico que se aplica a los enlaces electrónicos entre bodegas y almacenes, vendedores y compradores, etc., que permite intercambiar información sobre el estado del inventario en forma instantánea. Los sistemas de IED fueron comunes en algunas industrias durante años, y sólo ahora comienzan a incursionar en las demás industrias. Por ejemplo, los sistemas de intercambio al menudeo enlazan las tiendas con las bodegas centrales. La información de ventas se transmite en tiempo real usando lectores en los puntos de venta. Los gerentes pueden determinar al instante los niveles de inventario en la tienda. Los enlaces con las bodegas centrales eliminan la necesidad de levantamientos manuales de inventario y de decisiones de reabastecimiento. En los sistemas más avanzados, el intercambio electrónico de datos enlaza a los vendedores directamente con los clientes, para facilitar las compras JIT.

Aunque el IED se ha ganado su lugar en la administración de inventarios, todavía tiene un largo camino por recorrer. Se estima que unas 80 000 empresas estadounidenses con ganancias mayores a 5 o 6 millones de dólares usan el IED (Mohan, 1995). Esto representa sólo aproximadamente el 1 por ciento de los usuarios potenciales. Los que no lo usan temen que implementar esos sistemas cueste más que los posibles beneficios. También tienen en cuenta que los clientes prefieren sistemas basados en papeles.

En la implementación del intercambio electrónico de datos hay muchos éxitos documentados. Por ejemplo, La Strada Sportswear, fabricante de ropa deportiva para dama, embarca más de 50 000 prendas semanales y, a finales de los años ochenta, se vio forzada a entrar en el ámbito del IED por las demandas de los clientes. La empresa informa que este sistema facilita las comunicaciones, recorta los tiempos de procesamiento y agiliza las funciones de almacenaje y distribución (AIN, 1995). La industria de abarrotes se cuenta entre las que no han mostrado mayor entusiasmo en adoptar el IED. Una excepción es la cadena Giant Food, de 156 tiendas, que en fechas recientes desarrolló un sistema de suministro de tiendas con Nabisco y Frito-Lay manejado exclusivamente vía IED. El nuevo sistema abarca todas las fases del reabastecimiento de la tienda, incluyendo la colocación de pedidos y el pago de facturas (Garry, 1995).

Aplicación instantánea

LOS SISTEMAS DE INVENTARIO DE TRIAD SATISFACEN LAS NECESIDADES DEL MERCADO

Triad Systems Corporation, con sede en Livermore, California, desarrolla y vende una línea de productos para administrar información y controlar inventarios. La empresa da servicio a todos los ramos del mercado de refacciones automotrices, y también a otros mercados de mayoreo y menudeo. Los productos principales de Triad son sistemas que se venden a precio convencional. Esto es, se venden componentes y programas como un sistema integrado. Los sistemas de Triad, hasta este momento, se basan en el sistema UNIX con procesadores 486 y Pentium, y comprenden varios periféricos, como lectores en punto de venta, impresoras de etiquetas y de informes y lectores de códigos de barras.

Uno de sus clientes asiduos es Bill Clark, de Golden Big A Auto Parts en Golden, Colorado, quien afirma que con el sistema Triad puede revisar el estado de 72 000 partes en menos de 15 minutos por semana. Para él, una de las particularidades más valiosas es la capacidad que tiene el sistema para identificar los artículos "lentos", por lo que no tiene casos frecuentes de sobreabasto.

La meta de Triad, fundada a mediados de los años setenta, era encontrar una aplicación que usara el chip Z-80. (Esos chips fueron anteriores a los microprocesadores de las primeras computadoras personales, que para la PC de IBM eran los 8088.) Los fundadores de Triad pronto comenzaron a desarrollar sistemas de inventario para refaccionarias automotrices, que sigue siendo el principal nicho del mercado para la compañía.

Se podría suponer que un producto en un mercado especializado sólo tendría un potencial limitado de rentabilidad. Sin embargo, las ganancias de Triad han rebasado los 150 millones de dólares anuales desde 1992. La empresa augura un crecimiento mediante otras partes del mercado de refacciones automotrices (como, por ejemplo, los talleres de servicio); de hecho, aumentaron las ventas con otros mercados y lugares (particularmente en Europa). Han vendido sistemas a ferreterías, minoristas de decoración de casas y astilleros (patios de madera), así como a mayoristas automotrices.

La nueva generación de los sistemas Triad de control de inventario basados en UNIX usan muchas de las técnicas de pronóstico y control descritas en este libro. Entre ellas están el suavizamiento exponencial para rastrear el promedio y la variación de la demanda, el análisis ABC para identificar los artículos con mucho y poco movimiento, y cálculos de punto de reorden y dimensionamiento de lote, para obtener un costo mínimo y el nivel de servicio deseado.

Fuente: Reporte Anual 1994, literatura de promoción y entrevistas a empleados de Triad.

¿Qué revela la teoría acerca de las ventajas del IED para reducir los costos relacionados con inventarios? En primer lugar, el intercambio puede reducir los tiempos de demora de los pedidos. Las existencias de seguridad son directamente proporcionales al tiempo de demora del pedido, así que cuando se reducen los tiempos ocurre lo mismo con las existencias. En segundo lugar, que quizá sea el factor más importante, es que el intercambio debe reducir los costos fijos de pedido. Los tamaños de lote aumentan en proporción a la raíz cuadrada de los costos fijos, de modo que la disminución en éstos equivale a menores tamaños de lote. A su vez, los tamaños menores de lote dan como resultado pedidos más frecuentes, y en consecuencia se reducen los inventarios promedio. El resultado se parece más a un sistema del tipo JIT.

Deben atenderse asuntos legales y logísticos a medida que los sistemas de intercambio electrónico de datos ganan popularidad. Como hace notar Murray (1995), el código comercial uniforme pide que exista un contrato firmado disponible en la compra de bienes de 500 dólares o más. Esto significa que las órdenes de compra generadas electrónicamente por el intercambio de datos deben imprimirse y firmarse, eliminando así algunas de las ventajas del sistema sin documentos en papel. Si bien la transacción IED se almacena en un archivo electrónico, los tribunales deben determinar todavía si constituye un registro suficientemente apegado al código.

5.10 NOTAS HISTÓRICAS Y LECTURAS ADICIONALES

La investigación sobre modelos estocásticos de inventario fue impulsada por el esfuerzo bélico nacional y parece haber sido originada en la década de los cuarenta, aunque el primer material publicado apareció a principios de los años cincuenta (véase Arrow, Harris y Marschak, 1951;

Dvorestsky, Kiefer y Wolfowitz, 1952). El importante texto de Arrow, Karlin y Scarf (1958) causó gran interés en el campo.

El modelo descrito de un sistema de punto de reorden con tamaño de lote con criterio de desabasto parece atribuirse a Whitin (1957). Las extensiones a niveles de servicio se atribuyen a Brown (1967), quien también trata otros temas de importancia práctica, como la relación entre los modelos de pronóstico y el control de inventarios. Una excelente descripción de las curvas de intercambio y del análisis ABC y la presentación más profunda de algunos de los temas ya descritos puede encontrarse en el libro de Silver y Peterson (1985). Love (1979) presenta un excelente resumen de muchos de los temas tratados aquí. Hadley y Whitin (1963) también describen esos temas, con un nivel matemático más avanzado.

Se ha manifestado mucho interés en el desarrollo de políticas (s, S) aproximadas. Scarf (1960) y después Iglehart (1963) demostraron la optimalidad de esas políticas, y Veinott y Wagner (1965) consideraron métodos para calcular las políticas (s, S) óptimas. Se han sugerido varias técnicas de aproximación. Ehrhardt (1979) describe la aplicación del análisis de regresión para ajustar una red de políticas óptimas, mientras que Freeland y Porteus (1980) adaptan los métodos de programación dinámica al problema. Porteus (1985) compara numéricamente la eficacia de varias técnicas de aproximación. Kaplan (1970) y Nahmias (1979), entre otros, describen algunos de los asuntos concernientes a la incertidumbre del tiempo de demora.

La formulación original del problema típico de inventario de multiniveles provino de Clark y Scarf (1960). Sus extensiones han sido desarrolladas por varios investigadores, entre los cuales los más destacados son Bessler y Veinott (1966) y, en fecha más reciente, Federgruen y Zipkin (1984). Los sistemas de menudeo-bodega han sido estudiados por Deuermeyer y Schwarz (1981), y por Eppen y Schrage (1981). Schmidt y Nahmias (1985) analizaron un sistema de ensamble cuando la demanda del producto final es aleatoria.

El trabajo clásico sobre el modelo multiniveles que se ha constituido como la base de los sistemas de control de inventarios implementados entre los militares se debe a Sherbrooke (1968). Una extensión reciente de su análisis fue considerada por Graves (1985). Muckstadt y Thomas (1980) describen las ventajas de implementar modelos multiniveles en un ambiente industrial. Se puede encontrar una reseña detallada de los modelos para administrar artículos reparables en Nahmias (1981).

El interés en los modelos de control de inventario perecedero parece originarse en el problema del control en bancos de sangre, aunque tiene un impacto económico mucho mayor en la administración de productos alimenticios. La mayoría de los modelos matemáticos para artículos perecederos suponen que el inventario se despacha de las existencias en una base de primero el más antiguo (aunque una excepción notable es el estudio por Cohen y Pekelman, 1978). Nahmias (1982) presenta una reseña detallada del control de inventarios perecederos.

El problema de bienes estilísticos ha sido estudiado por varios investigadores. Entre los trabajos representativos en esta área se incluyen los de Murray y Silver (1966), de Hartung (1973), y de Hausman y Peterson (1972). Se recomienda a los interesados en lecturas adicionales sobre modelos estocásticos de inventarios los diversos artículos sobre el tema, entre los que destacan los de Scarf (1963), Veinott (1966) y Nahmias (1978). Una colección reciente de crónicas actualizadas se puede encontrar en Graves *et al.* (1992).

5.11 Resumen

Este capítulo presenta una perspectiva de varios métodos de control de inventarios cuando la demanda del artículo es aleatoria. El *modelo del repartidor de periódicos* se basa en la hipótesis de que el producto tiene una vida útil exactamente igual al periodo de planeación. Se supone que la demanda durante el periodo es una variable aleatoria continua con función de

distribución acumulada $F(x)$, y que hay costos especificados de sobrantes y faltantes, iguales a c_o y c_u que se cargan contra el inventario restante disponible al final del periodo, o al exceso de la demanda, respectivamente. La cantidad óptima de pedido Q^* resuelve la ecuación

$$F(Q^*) = \frac{c_u}{c_u + c_o}.$$

También se estudiaron la demanda discreta y varios periodos de planeación.

Desde un punto de vista práctico, el modelo del repartidor de periódicos tiene una limitación seria: no considera un costo positivo de colocación de pedido. Por esa razón se describe una extensión del modelo de CEP, llamado modelo de *tamaño del lote-punto de reorden.* La variable aleatoria clave en este caso fue la demanda durante el tiempo de la demora. Se indica cómo obtener valores óptimos de las variables de decisión Q y R con la solución iterativa de dos ecuaciones. El sistema funciona así: cuando el nivel de inventario disponible llega a R, se coloca un pedido por Q unidades (que llegarán transcurrido el tiempo de demora τ). Esta política es la base de muchos sistemas comerciales de control de inventario.

Con los niveles de servicio se dispone de una alternativa a los costos de faltantes. Se describieron dos niveles de servicio: Tipo 1 y Tipo 2. El Tipo 1 es la probabilidad de que no se agoten las existencias en un ciclo de pedido, y el Tipo 2 es la probabilidad de poder cumplir con una demanda cuando ocurre. A éste también se le llama *tasa de cumplimiento*, y es la definición más natural de servicio en la mayoría de las aplicaciones.

También se explican algunos temas adicionales relativos a los sistemas (Q, R). El *costo implícito de faltantes* es el costo efectivo de desabasto que resulta de la especificación del nivel de servicio. Suponiendo distribución normal, se indica cómo se transforma la distribución de la demanda periódica en la demanda durante el tiempo de demora. Por último, se consideran los efectos de la *variabilidad del tiempo de demora.*

Si se incluye un costo de preparación en la versión del problema del repartidor de periódicos para varios periodos, la forma óptima de la política de control se llama *política* (s, S). Esto significa que si el inventario inicial u en un periodo es menor o igual a s, se coloca un pedido por $S - u$ unidades. Una aproximación efectiva a la política (s, S) óptima puede obtenerse resolviendo el problema como si fuera revisión continua, para obtener la política (Q, R) correspondiente, e igualando $s = R$ y $S = Q + R$. También se describe la aplicación de niveles de servicio a los sistemas de revisión periódica.

La mayoría de los sistemas reales de inventario implican administrar más de un solo producto. Se describen varios tópicos que surgen al manejar un sistema de inventario con productos diversos. Uno de ellos es la cantidad de tiempo y los gastos que deben asignarse para controlar cada artículo. El *sistema ABC* es un método para clasificar los artículos de acuerdo con su volumen anual de ventas. Otro aspecto se refiere a la correcta elección de los parámetros de costo empleados para determinar las políticas de control de inventario. Como muchos de los parámetros de costo que se usan en el análisis de inventarios implican juicio gerencial y no se miden fácilmente, resultaría útil que la administración pudiera comparar los efectos de varios ajustes de parámetros y el desempeño del sistema. Una técnica conveniente para realizar tales comparaciones sería mediante las *curvas de intercambio.* Se describen dos de las más usadas: 1) el balanceo entre la inversión en inventarios y la frecuencia de reabastecimiento, y 2) el balanceo entre la inversión en existencias de seguridad y niveles de servicio.

El capítulo concluye con una descripción del intercambio electrónico de datos (IED), que consiste en enlazar electrónicamente distintas partes de un sistema logístico de, por ejemplo, bodegas y tiendas. Al reducir los costos de pedido y los tiempos de demora de pedidos, el intercambio puede significar niveles de inventario apreciablemente menores, sin pérdida de servicio. Algunas industrias, como la de ventas al menudeo, han abrazado este intercambio, mientras que otras, como las alimentarias, adoptan lentamente las nuevas tecnologías.

Problemas adicionales sobre modelos estocásticos de inventario

28. Una tienda de material para artistas tiene distintos artículos en existencia para satisfacer las necesidades de los pintores. En cada caso de los descritos a continuación, ¿cuál es el modelo adecuado de control de inventarios que debe adoptar esa tienda para administrar el reabastecimiento del artículo en cuestión? Escoja entre las propuestas de la siguiente lista, y explique su respuesta en cada caso:

CEP simple	Repartidor de periódicos con nivel de servicio
Tasa finita de producción	Modelo (Q, R) con costo de faltantes
CEP con descuentos por cantidad	Modelo (Q, R) con nivel de servicio Tipo 1
CEP restringida por recursos	Modelo (Q, R) con nivel de servicio Tipo 2
Modelo del repartidor de periódicos	Otro tipo de modelos

a) Un diluyente muy volátil se pide cada tres meses. Las latas que no se venden durante el periodo de tres meses se desechan. La demanda varía mucho de un trimestre al otro.

b) Una pintura blanca al óleo se vende con una tasa bastante estable de 600 tubos por mes, y necesita un tiempo de demora de pedido de seis semanas. La tienda la compra a 1.20 dólares por tubo.

c) La pintura al óleo de siena no se vende tan uniformemente como la blanca. Las ventas de siena varían mucho de un mes al siguiente. La vida útil de la pintura es dos años, aproximadamente, pero la tienda vende casi toda la siena antes del límite de dos años.

d) Los pinceles de cerdas sintéticas se compran a un proveedor en la Costa Este, quien cobra 1.60 dólares por cada pincel en pedidos de menos de 100, y 1.30 dólares por pieza en pedidos de 100 o mayores. La tienda vende los pinceles con una frecuencia bastante uniforme de 40 al mes, a 2.80 dólares cada uno.

e) Los pinceles con cerdas de pelo de camello se compran al mismo proveedor, que ofrece un precio de descuento similar al de los de cerdas sintéticas. Sin embargo, los pinceles de camello muestran variaciones considerables en la ventas de cada mes.

29. Cada año la demanda de lápices #2 en la papelería de la universidad tiene distribución normal con media 1 000 y desviación estándar 250. La papelería compra los lápices a .6 y los vende a .20 cada uno. Hay un tiempo de demora de dos meses desde el inicio hasta el recibo del pedido. El contador estima que el costo es 20 dólares por pedido, teniendo en cuenta el tiempo del empleado al llevar a cabo los trámites necesarios para iniciar y recibir un pedido, y recomienda una tasa anual de interés de 22 por ciento para determinar el cargo por inventarios. El costo de faltantes es el de la ganancia perdida más .20 adicionales por lápiz, que representan el costo de pérdida de buena voluntad.

a) Calcule el valor óptimo del punto de reorden R, suponiendo que el tamaño de lote es igual a la cantidad económica de pedido.

b) Calcule los valores óptimos simultáneos de Q y R.

c) Compare los costos anuales promedio de inventario, preparación y desabasto para las políticas determinadas en las partes *a*) y *b*).

d) ¿Cuál es la existencia de seguridad para este artículo en la solución óptima?

30. Para el problema de satisfacer la demanda de lápices #2 en la papelería universitaria del problema 29:

a) Resuelva el problema sustituyendo el costo de faltantes con un criterio de servicio Tipo 1 igual a 95 por ciento.

b) Resuelva el problema sustituyendo el costo de faltantes con un criterio de servicio Tipo 2 igual a 95 por ciento. Suponga que Q es la cantidad más económica de pedido (CEP).

c) Calcule los valores óptimos simultáneos de Q y R, suponiendo un nivel de servicio Tipo 2 igual a 95 por ciento.

31. Conteste si lo siguiente es cierto o falso.

a) El tiempo de demora siempre es menor que el tiempo de ciclo.

b) El tamaño óptimo de lote para un objetivo de servicio Tipo 1 igual a X por ciento siempre es menor que el tamaño óptimo de lote para un objetivo de servicio Tipo 2 igual a X por ciento para el mismo artículo.

c) El modelo del repartidor de periódicos no incluye un costo fijo de pedido.

d) El análisis ABC clasifica a los artículos según el valor anual de su demanda.

e) Para un modelo de tasa finita de producción, el tamaño óptimo de lote por producir cada ciclo es igual al inventario máximo en cada ciclo.

32. Uno de los productos que maneja la tienda de pinturas Weiss del problema 14 es un solvente muy volátil que, debido a cambios químicos del producto, tiene una vida en almacén de exactamente un año. Weiss lo adquiere a 20 dólares por lata de 1 galón y lo vende a 50 dólares por lata. El proveedor vuelve a comprar las latas que no se venden durante el año, a 8 dólares, para reprocesarlas. La demanda de este solvente varía generalmente de 20 a 70 latas por año. Weiss asigna un cargo de inventario a las latas no vendidas con base en 30 por ciento de tasa anual de interés.

a) Suponiendo que son igualmente probables todos los valores de la demanda de 20 a 70, ¿cuál es la cantidad óptima de latas de solvente que debe comprar Weiss cada año?

b) Un análisis más exacto de la demanda indica que una distribución normal se ajusta mejor a los datos. La media de la distribución es idéntica a la que se usó en la parte *a*), y el estimador de desviación estándar resulta ser 7. ¿Qué política obtiene usted ahora?

33. Semicon es una nueva empresa que produce semiconductores para varias aplicaciones. El proceso de grabar los circuitos requiere grandes cantidades de ácido nítrico, cuya vida en almacén sólo es de tres meses. Semicon estima que necesitará de 1 000 a 3 000 galones del ácido para el siguiente periodo trimestral, y supone que todos los valores en este intervalo tienen la misma probabilidad. El ácido les cuesta 150 dólares por galón, y se supone una tasa de interés anual de 30 por ciento para la inversión en inventario, y que el almacenamiento del ácido cuesta 35 dólares por galón. (Suponga que todos los costos de inventario se asignan al final del periodo trimestral.) Cuesta 75 dólares por galón desechar el ácido al final del periodo. Si a la empresa se le agota ese ácido durante el periodo trimestral, puede comprar suministro de emergencia a 600 dólares por galón.

a) ¿Cuántos galones de ácido nítrico debe comprar Semicon? La experiencia posterior en el mercado indica que la demanda se parece más a una distribución normal, con media 1 800 y desviación estándar 480.

b) Suponga que ahora Semicon cambia a un criterio de tasa de cumplimiento de 94 por ciento. ¿Cuántos galones debe comprar ahora, al iniciar cada periodo trimestral?



34. *Simulador del repartidor de periódicos.* Para resolver este problema su hoja de cálculo debe tener una función que produzca números aleatorios [@RAND en Lotus 1-2-3, ALEATORIO( ) en Excel en español, y RAND( ) en Excel en inglés]. El objetivo de este ejercicio es construir un sistema de simulación de inventario con revisión periódica con demanda aleatoria. Suponemos que el lector está familiarizado con la simulación de Monte Carlo.

Su hoja de cálculo debe permitir lugares de celda para almacenar valores del costo de inventario, costo de penalización, costo proporcional al pedido, cantidad de pedido hasta el punto, inventario inicial y para promedio, y desviación estándar de la demanda periódica.

Un método eficiente de generar una observación a partir de una variable normal estándar es la fórmula

$$Z = [-2\ln(U_1)]^{0.5}\cos(2\pi U_2),$$

en donde U_1 y U_2 son dos elecciones distintas de una distribución uniforme (0, 1). (Véase, por ejemplo, Fishman, 1973.) Observe que se necesitan dos llamadas independientes de @RAND. Como Z es aproximadamente normal estándar, la demanda X está dada por

$$X = \sigma Z + \mu$$

en donde μ y σ son la media y la desviación estándar de la demanda en un periodo.

Una distribución de la hoja de cálculo podría ser la siguiente:

SIMULADOR DE REPARTIDOR DE PERIÓDICOS

Costo de inventario =	**Media de la demanda =**
Costo de pedido =	**Desviación estándar de la demanda =**
Costo de penalización =	**Inventario inicial =**
	Pedir hasta el punto =

Periodo	Inventario inicial	Cantidad de pedido	Demanda	Inventario final	Costo de Inventario	Costo de penalización	Costo de pedido
1							
2							
3							
.							
.							
.							
20							
Totales							

Cada vez que vuelva a realizar el cálculo, el simulador generará una sucesión distinta de demandas. Un conjunto de parámetros podría ser el siguiente:

$$h = 2,$$
$$c = 5,$$
$$p = 20,$$
$$\mu = 100,$$
$$\sigma = 20,$$
$$I_0 = 50,$$
$$\text{Pedir hasta el punto} = 150.$$



35. *Uso del simulador para optimización.* Este problema se puede resolver sólo si el programa que usa no restringe el tamaño de la hoja de cálculo. Suponga los parámetros del problema 34, pero use $\sigma = 10$. Amplíe la hoja de cálculo del problema 34 a 1000 renglones o más, y calcule el costo promedio por periodo. Si el costo cambia mucho al recalcular, podría necesitar más renglones.

    Ahora haga la prueba con distintos valores del parámetro "pedir hasta el punto", para determinar cuál es el que minimiza el costo promedio por periodo. Compare sus resultados con la solución teórica óptima.



36. *Calculadora del repartidor de periódicos.* Diseñe una hoja de cálculo para determinar el punto óptimo hasta el cual pedir, para un problema de repartidor de periódicos con demandas normalmente distribuidas. Para evitar búsquedas en la tabla, use la fórmula de aproximación

$$Z = 5.0633[F^{0.135} - (1 - F)^{0.135}]$$

para la inversa de la función de distribución normal estándar. (Esta fórmula se tomó de Ramberg y Schmieser, 1972.) El pedido óptimo hasta el punto, y^*, tiene la forma $y^* = \sigma Z + \mu$.

a) Suponga que todos los parámetros son los del problema 34. Grafique y^* como función de p, el costo de penalización, para $p = 10, 15, \ldots, 100$. ¿En qué porcentaje aumenta y^* si p aumenta de 10 a 20? ¿Y de 50 a 100?

b) Repita la parte *a*) con $\sigma = 35$. Comente el efecto que tiene la varianza sobre la sensibilidad de y^* respecto a p.

37. Un gran productor nacional de alimentos enlatados desea comprar 100 cosechadoras que puedan adaptarse a sus necesidades. Una de las partes usadas en la cosechadora es una cuchilla reemplazable para la cosecha de maíz. Las cuchillas de repuesto pueden comprarse cuando se coloca el pedido a 100 dólares cada una, pero si se compran después, costarán 1000 dólares cada una debido a que se requerirá una corrida especial de producción.

    Se estima que la cantidad de hojas de reemplazo que requiere una trilladora en su vida útil puede aproximarse muy bien con una distribución normal con media de 18 hojas, y desviación estándar 5.2. El fabricante de la cosechadora está de acuerdo en volver a comprar las hojas no usadas a 20 dólares cada una. ¿Cuántas cuchillas de repuesto debe adquirir la empresa junto con las cosechadoras?

38. Crazy Charlie's, una tienda de descuento que vende equipos estereofónicos, usa suavizamiento exponencial simple, con constante de suavizamiento $\alpha = .2$, para rastrear la media y la desviación absoluta media de la demanda mensual de sus artículos. Uno de ellos, un receptor estereofónico, ha tenido las siguientes ventas durante los últimos tres meses: 126, 138 y 94.

    Hace tres meses, la computadora registraba los valores de media = 135 y DAM = 18.5.

a) Aplique las ecuaciones de suavizamiento exponencial de la sección 5.1 para calcular los valores actuales de la media y la DAM de la demanda mensual. (Suponga que los valores almacenados fueron calculados *antes* de observar la demanda de 126.)

b) Suponga que el tiempo de demora de pedido para este artículo particular es 10 semanas (2.5 meses). Determine los estimados actuales de la media y la desviación estándar de la demanda durante el tiempo de demora suponiendo una distribución normal mensual.

c) Este radio en particular se compra directamente en Japón, y por ello ha habido mucha variación en el tiempo de demora del reabastecimiento de un ciclo al siguiente. Con base en un análisis de ciclos anteriores de pedido, se estima que la desviación estándar del tiempo de demora es 3.3 semanas. [Las demás cantidades relevantes son las de las partes *a*) y *b*)]. Calcule la media y la desviación estándar de la demanda en tiempo de demora para este caso.

d) Si Crazy Charlie's usa un objetivo de servicio Tipo 1 igual a 98 por ciento para controlar el reabasto de este artículo, ¿cuál es el valor del nivel de reorden? Suponga que la demanda en el tiempo de demora tiene distribución normal.

39. El departamento de electrodomésticos de una gran tienda usa un sistema tamaño del lote-punto de reorden para controlar el reabastecimiento de cierto modelo de radio de mesa FM. Se vende un promedio de 10 radios cada semana. La demanda semanal se apega a una distribución normal con varianza 26.

    La tienda paga 20 dólares por cada radio, que vende en 75 dólares. Los costos fijos de reabastecimiento ascienden a 28 dólares. El departamento de contabilidad recomienda una tasa de interés de 20 por ciento para el costo del capital. Los costos de almacenamiento son de 3 por ciento, y de averías, 2 por ciento del valor de cada radio.

    Si un cliente pide el radio cuando no lo hay, por lo general irá a otra parte. Se estima que los costos de pérdida de buena voluntad son unos 25 dólares por radio. El tiempo de demora del pedido es tres meses.

    a) Si los tamaños de lote se basan en la fórmula de CEP, ¿qué nivel de reorden se debe usar con estos radios?

    b) Calcule los valores óptimos de (Q, R).

    c) Compare los costos promedio anuales de inventario, pedido y desabasto para las políticas que determinó en las partes *a*) y *b*).

    d) Resuelva de nuevo el problema con las ecuaciones (1) y (2´), y no con las (1) y (2). ¿Cuál es el efecto de incluir en forma explícita las ventas perdidas?

40. Vuelva a resolver el problema de la tienda de departamentos descrito en el problema 39, reemplazando el costo de faltantes con un nivel de servicio Tipo 1 de 96 por ciento.

41. Resuelva nuevamente el problema 39, reemplazando el costo de faltantes con un nivel de servicio Tipo 2 igual a 96 por ciento. ¿Cuál es el costo de faltantes implícito?

42. Vea la ecuación para determinar el costo anual promedio esperado de la política (Q, R) en un sistema de control de inventario con revisión continua, en la sección 4:

$$G(Q, R) = h\left(\frac{Q}{2} + r - \lambda T\right) + \frac{K\lambda}{Q} + \frac{P\lambda n(R)}{Q}.$$

Diseñe una hoja de cálculo para calcular $G(Q, R)$ para un intervalo de valores de $Q \geq$ CEP y $R \geq \mu$. Use la siguiente fórmula de aproximación para $L(z)$ para evitar búsquedas en tabla:

$$L(z) = \exp(-0.92 - 1.19z - 0.37z^2).$$

(Esta fórmula es de Parr, 1972.) Guarde los parámetros c, h, p, μ, σ y λ en ubicaciones de celda. Busque visualmente entre los valores tabulados de $G(Q, R)$ para determinar el valor mínimo y estimar los valores óptimos de (Q, R) en esta forma. Compare sus resultados con el óptimo verdadero que determinó con cálculos manuales.

a) Resuelva el ejemplo 5.4.

b) Resuelva el problema 13 con este método.

c) Resuelva el problema 14 con este método.

41. La demanda diaria de una parte de repuesto de motor es una variable aleatoria con la distribución siguiente, basada en la experiencia:

Número de demandas por día	Probabilidad
0	.21
1	.38
2	.19
3	.14
4	.08

Se espera que la parte sea obsoleta al cabo de 400 días. Suponga que las demandas de un día al siguiente son independientes. Esa parte cuesta 1 500 dólares por pieza cuando se adquiere por adelantado al periodo de 400 días, y 5 000 dólares por pieza cuando se compran de emergencia, durante el periodo de 400 días. Los costos de inventario de las partes no usadas se basan en una tasa diaria de interés igual a .08 por ciento. Las partes no usadas se pueden vender a 10 por ciento de su precio de compra. ¿Cuántas partes deben adquirirse por adelantado al periodo de 400 días? (Sugerencia: Sean $D_1, D_2, \ldots, D_{400}$ las demandas diarias de la parte. Suponga que cada D_1 tiene una media μ y una varianza σ^2. El teorema del límite central establece que la demanda total del periodo de 400, que es ΣD_i, tiene distribución aproximadamente normal con media 400μ y varianza $400\sigma^2$.)

44. Cassorla's Clothes vende una gran cantidad de camisas blancas de vestir. Éstas llevan la etiqueta de su tienda, las embarca un fabricante en Nueva York. Cassorla, el propietario, dice: "Deseo asegurarme de nunca tener faltantes de camisas. Siempre trato de mantener una existencia mínima para dos meses. Cuando mi inventario baja de ese nivel, coloco otro pedido por dos meses de suministro. Durante 20 años he estado usando ese método, y funciona bien".

Las camisas le cuestan 6 dólares y las vende a 15 dólares cada una. El costo de procesar un pedido y recibirlo asciende a 80 dólares, y se necesitan tres semanas para recibir el embarque. La demanda mensual tiene una distribución aproximadamente normal, con media 120 y desviación estándar 32. Para calcular el costo de mantener el inventario, suponga una tasa anual de interés de 20 por ciento.

a) ¿Qué valores de Q y R está utilizando Hy Cassorla para controlar el inventario de camisas blancas de vestir?

b) ¿Qué tasa de cumplimiento (nivel de servicio Tipo 2) se alcanza con la política actual?

c) Con base en un criterio de tasa de cumplimiento igual a 99 por ciento, calcule los valores óptimos de Q y R que debe usar. Suponga en los cálculos que un mes equivale a cuatro semanas.

d) Calcule la diferencia en los costos promedio anuales de inventario y preparación entre las políticas *b*) y *c*).

e) Estime cuánto tiempo se necesitaría para pagar un sistema de control de inventario que cuesta 25 000 dólares, suponiendo que las camisas de vestir representan el 5 por ciento de los negocios anuales de Cassorla, y que se pueden realizar ahorros semejantes también en los demás artículos.

Apéndice 5-A

Convenciones de notación y repaso de probabilidad

La demanda se representará por D, y se supone que es una variable aleatoria. La función de distribución acumulada de la demanda es $F(x)$, definida por

$$F(x) = P\{D \leq x\} \qquad \text{para } -\infty < x < +\infty.$$

Cuando D es continua, la función de densidad de probabilidades de la *demanda*, $f(x)$, se define como sigue:

$$f(x) = \frac{dF(x)}{dx}.$$

Cuando D es discreta, $f(x)$ es la función de probabilidad (fp), definida por

$$f(x) = P\{X = x\} = F(x) - F(x-1).$$

Observe que en el caso continuo, la función de demanda no es una probabilidad, y que el valor de $f(x)$ no necesariamente es menor que 1, aunque siempre se cumple que $f(x) \geq 0$ para toda x.

El valor esperado de la demanda, $E(D)$, se define como sigue:

$$E(D) = \int_{-\infty}^{+\infty} xf(x)\,dx$$

en el caso continuo, y

$$E(D) = \sum_{x=-\infty}^{+\infty} xf(x)$$

en el caso discreto.

Usaremos el símbolo μ para representar el valor esperado de la demanda $[E(D) = \mu]$. En lo que sigue supondremos que D es continua; hay fórmulas parecidas que son válidas en el caso discreto. Sea $g(x)$ cualquier función de valor real de la variable real x. Entonces

$$E(g(D)) = \int_{-\infty}^{+\infty} g(x)f(x)\,dx.$$

En particular, sea $g(D) = \text{máx}(0, Q - D)$. Entonces

$$E(g(D)) = \int_{-\infty}^{+\infty} \text{máx}\,(0, Q - x)f(x)\,dx.$$

Como la demanda anual es no negativa, se debe cumplir que $f(x) = 0$ cuando $x < 0$. Además, cuando $x > Q$, $\text{máx}(0, Q - x) = 0$, así que se puede escribir

$$E(g(D)) = \int_{0}^{Q} (Q - x)f(x)\,dx.$$

En el análisis del modelo del repartidor de periódicos se aplica la regla de Leibniz para determinar la derivada de $G(Q)$. Según esa regla

$$\frac{d}{dy}\int_{a_1(y)}^{a_2(y)} h(x, y)\,dx = \int_{a_1(y)}^{a_2(y)} [\partial h(x, y)/\partial y]\,dx + h(a_2(y), y)a_2'(y) - h(a_1(y), y)a_1'(y).$$

Apéndice 5-B

Resultados adicionales y extensiones para el modelo del repartidor de periódicos

1. INTERPRETACIÓN DE LOS COSTOS DE EXCESO Y FALTANTES PARA EL PROBLEMA DEL PERIODO ÚNICO

Se definen

S = Precio de venta del artículo.

c = Costo variable del artículo.

h = Costo de mantener el inventario por unidad que permanece en almacén al final del periodo.

p = Costo de pérdida de buena voluntad más gastos contables (que se cargan contra la cantidad de pedidos pospuestos en los libros, al final del periodo).

Indicaremos cómo deben interpretarse c_u y c_o en términos de esos parámetros. Como antes, sean Q la cantidad pedida y D la demanda durante el periodo. Suponga, sin pérdida de generalidad, que el inventario inicial es cero. Entonces, el costo incurrido al final del periodo es

$$cQ + h\,\text{máx}(Q - D, 0) + p\,\text{máx}(D - Q, 0) - S\,\text{mín}(Q, D).$$

El costo esperado es

$$G(Q) = cQ + h\int_0^Q (Q - x)f(x)\,dx + p\int_Q^\infty (x - Q)f(x)\,dx$$
$$- S\int_0^Q xf(x)\,dx - SQ\int_Q^\infty f(x)\,dx.$$

Usando

$$\int_0^Q xf(x)\,dx = \int_0^\infty xf(x)\,dx - \int_Q^\infty xf(x)\,dx = \mu - \int_Q^\infty xf(x)\,dx,$$

el costo esperado se puede escribir como sigue:

$$G(Q) = cQ + h\int_0^Q (Q - x)f(x)\,dx + (p + S)\int_Q^\infty (x - Q)f(x)\,dx - S\mu.$$

La cantidad óptima de pedido satisface a

$$G'(Q) = 0$$

es decir

$$c + hF(Q) - (p + S)(1 - F(Q)) = 0,$$

que da como resultado

$$F(Q) = \frac{P + S - c}{p + S + h}.$$

Al igualar $c_u = p + S - c$ y $c_o = h + c$ se obtiene la relación crítica en la forma $c_u/(c_u + c_o)$.

2. EXTENSIÓN A HORIZONTE INFINITO SUPONIENDO ACUMULACIÓN COMPLETA DE LA DEMANDA

Sea $D_1, D_2, \ldots$ una sucesión infinita de demandas. Suponga que las demandas son variables aleatorias independientes e idénticamente distribuidas, que tienen la función común de distribución $F(x)$ con la densidad $f(x)$. La política es pedir hasta Q cada periodo. Como todo el exceso de demanda se acumula, la cantidad del pedido en cualquier periodo es exactamente igual a la demanda del periodo anterior. La cantidad de unidades vendidas también será igual a la demanda. Para visualizarlo, consideremos la cantidad de unidades vendidas en periodos sucesivos:

$$\text{Cantidad de unidades vendidas en el periodo } 1 = \text{mín}(Q, D_1),$$

$$\text{Cantidad de unidades vendidas en el periodo } 2 = \text{máx}(D_1 - Q, 0) + \text{mín}(Q, D_2).$$

$$\text{Cantidad de unidades vendidas en el periodo } 3 = \text{máx}(D_2 - Q, 0) + \text{mín}(Q, D_3),$$

y así sucesivamente.

Estas relaciones son consecuencia de que acumular el exceso de demanda equivale a que las ventas se hagan en el periodo siguiente. Ahora, observe que

$$\text{mín}(Q, D_i) + \text{máx}(D_i - Q, 0) = D_i \qquad \text{para } i = 1, 2, \ldots,$$

que se obtiene al considerar los casos $Q < D_i$ y $Q \geq D_i$.

Por consiguiente, el costo esperado durante n periodos es

$$\begin{aligned} &cQ + (c - S)E(D_1 + D_2 + \cdots + D_{n-1}) - (S)E[\text{mín}(Q, D_n)] + nL(Q) \\ &\quad = cQ + (c - S)(n - 1)\mu - (S)E[\text{mín}(Q, D_n)] + nL(Q) \end{aligned}$$

donde

$$L(Q) = h\int_0^Q (Q - x)f(x)\,dx + p\int_Q^\infty (x - Q)f(x)\,dx.$$

Al dividir entre n y hacer que $n \to \infty$ se obtiene el promedio del costo para una cantidad infinita de periodos, que es

$$(c - S)\mu + L(Q).$$

El valor óptimo de Q se presenta cuando $L'(Q) = 0$, y así se obtiene

$$F(Q) = p/(p + h).$$

3. EXTENSIÓN A UN HORIZONTE INFINITO SUPONIENDO VENTAS PERDIDAS

Si el exceso de demanda se pierde, en lugar de acumularse, el argumento anterior ya no es válido. La cantidad de unidades vendidas en el periodo 1 es $\text{mín}(Q, D_1)$, que también es la cantidad de unidades vendidas en el periodo 2; la cantidad de unidades vendidas en el periodo 2 es $\text{mín}(Q, D_2)$ (ya que no hay acumulación del exceso de demanda), que también es la cantidad de unidades vendidas en el periodo 3, y así sucesivamente. Como se demostró en la sección 1 de este apéndice,

$$E[\text{mín}(Q, D)] = \mu - \int_Q^\infty (x - Q)f(x)\,dx.$$

En consecuencia, el costo esperado durante n periodos se expresa como sigue:

$$cQ + [(n-1)c - nS]\left[\mu - \int_Q^{\infty}(x-Q)f(x)\,dx\right] + nL(Q).$$

Si dividimos entre n y hacemos que $n \to \infty$, obtenemos la siguiente expresión del costo promedio por periodo:

$$(c-S)\left[\mu - \int_Q^{\infty}(x-Q)f(x)\,dx\right] + L(Q).$$

Al diferenciar con respecto a Q e igualar el resultado a cero se obtiene la condición siguiente para la Q óptima:

$$F(Q) = \frac{p+S-c}{p+S+h-c}.$$

Si se igualan $c_u = p + S - c$ y $c_o = h$ se obtiene la relación crítica en la forma $c_u/(c_u + c_o)$. Por consiguiente, interpretaremos a c_u como el costo de pérdida de buena voluntad más la utilidad perdida por venta, y a c_o sólo como el costo de inventario.

Apéndice 5-C

Deducción de la política (Q, R) óptima

Según la sección 5.4, el objetivo es calcular los valores de las variables Q y R que minimicen la función

$$G(Q, R) = h(Q/2 + R - \lambda\tau) + K\lambda/Q + p\lambda n(R)/Q. \tag{1}$$

Como esta función se minimizará con respecto a las dos variables (Q, R), una condición necesaria para la optimalidad es que $\partial G/\partial Q = \partial G/\partial R = 0$. Las dos ecuaciones que resultan son

$$\frac{\partial G}{\partial Q} = \frac{h}{2} - \frac{K\lambda}{Q^2} - \frac{p\lambda n(R)}{Q^2} = 0, \tag{2}$$

$$\frac{\partial G}{\partial R} = h + p\lambda n'(R)/Q = 0. \tag{3}$$

Observe que ya que $n(R) = \int_R^{\infty}(x-R)f(x)\,dx$, se puede demostrar que

$$n'(R) = -(1 - F(R)).$$

Partiendo de la ecuación (2) obtenemos

$$\frac{1}{Q^2}[K\lambda + p\lambda n(R)] = \frac{h}{2}$$

o bien

$$Q^2 = \frac{2K\lambda + 2p\lambda n(R)}{h},$$

que da como resultado

$$Q = \sqrt{\frac{2\lambda[K + pn(R)]}{h}}. \tag{4}$$

A partir de la ecuación (3) obtenemos

$$h + p\lambda[-(1 - F(R))]/Q = 0,$$

que da como resultado

$$1 - F(R) = Qh/p\lambda. \tag{5}$$

Apéndice 5-D

Distribuciones de probabilidad para administración de inventarios

En este capítulo nos hemos referido con frecuencia a la distribución normal como modelo de incertidumbre de la demanda. Aunque es cierto que domina en las aplicaciones, no es la única alternativa disponible. De hecho podría ser una mala opción en ciertos casos. En este apéndice describiremos otras distribuciones para representar la incertidumbre de la demanda.

1. LA DISTRIBUCIÓN DE POISSON COMO MODELO DE INCERTIDUMBRE DE LA DEMANDA

Cuando los artículos tienen poco movimiento, esto es, que tienen tasas de demanda pequeñas, el modelo de distribución normal no es el adecuado. Como éste tiene sus extremos en el infinito, cuando la media es pequeña es posible que una gran parte de la curva de densidad se prolongue en el semiplano de las x negativas. Por ello se podrían obtener malos resultados en los cálculos de existencias de seguridad. Una alternativa común para representar los artículos con poco movimiento es la distribución de Poisson, que es discreta y se define sólo en el semiplano de las x positivas. Sea X la distribución de Poisson con parámetro μ. Entonces

$$f(x) = \frac{e^{-\mu}\mu^x}{x!} \qquad \text{para } x = 0, 1, 2, \ldots .$$

(La deducción de la distribución de Poisson y su relación con la distribución exponencial se describen con detalle en la sección 12.3.)

Una propiedad importante de la distribución de Poisson es que tanto la media como la varianza son iguales a μ (y entonces $\sigma = \sqrt{\mu}$). Por consiguiente, la desviación estándar observada de la demanda periódica (o la demanda en el tiempo de demora) debería aproximarse a la raíz cuadrada de la demanda periódica media (o la demanda en tiempo de demora), para que la distribución sea adecuada. La tabla A-3, al final del libro, corresponde a la distribución de Poisson complementaria acumulada. Esta tabla permite calcular las políticas óptimas para los modelos del repartidor de periódicos y (Q, R), suponiendo que la demanda sigue dicha distribución.

Indicaremos cómo determinar las políticas óptimas para los modelos del repartidor de periódicos y de (Q, R) cuando la demanda sigue la distribución de Poisson. Para el caso del repartidor de periódicos, tan sólo se aplica el método descrito en la sección 5.3 para demanda discreta, y se obtienen las probabilidades en la tabla A-3. Ilustraremos esto con un ejemplo.

Ejemplo *5D.1* Regresemos al puesto de periódicos de Mac, descrito en varios ejemplos del capítulo 5. Algunos de sus clientes regulares le han pedido apartar determinada revista de equipos de audio. Mac está de acuerdo, aunque las ventas de esa revista han estado bajas. Con base en los datos anteriores, ha encontrado que

una distribución con media 5 se ajusta bien al patrón semanal de ventas. Mac compra las revistas a .50 dólares y las vende a 1.50 dólares. Regresa los ejemplares no vendidos a su proveedor, que se las paga a .10 dólares cada una. Calcule la cantidad de revistas que debe comprar cada semana a su proveedor.

Solución

El costo de tener excedentes es $c_o = .50 - .10 = .40$ dólares, y el de tener faltantes es $c_u = 1.50 - .50 = 1.00$ dólar. En consecuencia, la relación crítica es $c_u/(c_o + c_u) = 1/1.4 = .7143$. Como la tabla A-3 proporciona la distribución acumulada complementaria, hay que restar de 1 los elementos de esa tabla para obtener los valores de la función de distribución acumulada. Vemos en la tabla que

$$F(5) = 1 - .5595 = .4405,$$
$$F(6) = 1 - .3840 = .6160,$$
$$F(7) = 1 - .2378 = .7622.$$

Como la relación crítica está entre $F(6)$ y $F(7)$, pasamos al valor mayor y así obtenemos la cantidad óptima de pedido de 7 revistas.

No es difícil calcular las políticas (Q, R) óptimas cuando la demanda sigue una distribución de Poisson, pero para ello se requiere una solución iterativa parecida a la de la distribución normal descrita en la sección 5.5 de este capítulo. Se define $P(x)$ como el acumulado complementario de la distribución de Poisson. Esto es,

$$P(x) = \sum_{k=x}^{\infty} f(k).$$

La tabla A-3 muestra los valores de $P(x)$. Se puede demostrar entonces que, para la distribución de Poisson

$$n(R) = \mu P(R) - RP(R + 1),$$

y eso significa que toda la información necesaria para calcular las políticas óptimas aparece en la tabla A-3.[1] Es decir, en este caso no se requieren tablas especiales para calcular $n(R)$. Esta relación permite calcular la política (Q, R) usando el par de ecuaciones (1) y (2) de la sección 5.4.

Ejemplo *5D.2*

Una tienda de departamentos usa una política (Q, R) para controlar sus inventarios. Las ventas de un radio FM de bolsillo son en promedio de 1.4 a la semana. El radio le cuesta 50 dólares a la tienda, y los costos fijos de reabasto ascienden a 20 dólares. Los costos de inventario se basan en una tasa de interés anual de 20 por ciento, y el gerente de la tienda estima un costo de 12.50 dólares por venta perdida y pérdida de buena voluntad, si se le pide un radio cuando no lo tiene en existencia. El tiempo de demora del pedido es dos semanas, y con estudios estadísticos se ha demostrado que la demanda durante el tiempo de demora se describe muy bien con una distribución de Poisson. Calculemos los tamaños óptimos de lote y sus puntos de reorden correspondiente.

Solución

Los parámetros importantes para este problema son

$$\lambda = (1.4)(52) = 72.8.$$
$$h = (50)(0.2) = 10.0.$$
$$K = \$20.$$
$$\mu = (1.4)(2) = 2.8.$$
$$p = \$12.50.$$

Para iniciar el proceso de solución calculamos la cantidad económica de pedido. Es

$$\text{CEP} = \sqrt{\frac{2K\lambda}{h}} = \sqrt{\frac{(2)(20)(72.8)}{10}} = 17.1.$$

El siguiente paso es determinar R resolviendo

$$P(R) = Qh/p\lambda.$$

[1] Véase Hadley y Whitin (1963), p. 441.

Al sustituir Q con la CEP y hacer las operaciones se obtiene $P(R_0) = .1868$. Vemos en la tabla A-3 que $R = 4$ da como resultado $P(R) = .3801$, y que $R = 5$ produce $P(R) = .1523$. Suponiendo que se sigue la estrategia conservadora de redondear a la R más grande, seleccionaríamos $R = 5$ y $P(R) = .1523$. Por lo anterior,

$$n(R_0) = (2.8)(.1523) - (5)(.0651) = 0.1009.$$

Por consiguiente, Q_1 es

$$Q_1 = \sqrt{\frac{(2)(72.8)[20 + (12.5)(0.1009)]}{10}} = 17.6.$$

De aquí, $P(R_1) = .1934$, dando $R_1 = R_0 = 5$. Entonces, la solución ha convergido, porque los valores sucesivos de R (y por consiguiente de Q) son iguales. La solución óptima es $(Q, R) = (18, 5)$.

2. LA DISTRIBUCIÓN DE LAPLACE

La distribución de Laplace ha sido sugerida para modelar artículos con poco movimiento o con más varianza en sus extremos que la distribución normal. Se le ha llamado pseudoexponencial, porque matemáticamente es una distribución exponencial con una imagen especular simétrica. (Véase una explicación detallada sobre ésta en el capítulo 12, en el contexto de la administración de confiabilidad.)

La forma matemática de la función de densidad de probabilidad de Laplace es

$$f(x) = \frac{1}{2\theta}\exp(-|x - \mu|/\theta) \qquad \text{para } -\infty < x < +\infty.$$

Como la función de distribución de probabilidad es simétrica respecto a μ, la media es μ. La varianza es $2\theta^2$. La distribución de Laplace también es un modelo razonable para artículos con poco movimiento, y es una alternativa de la distribución normal para artículos con mucho movimiento cuando hay más dispersión en los extremos de la distribución que los correspondientes a la distribución normal. Presutti y Trepp (1970) observaron que simplificaba mucho el cálculo de la política óptima para el modelo (Q, R).

Se puede demostrar que para cualquier valor de $R > \mu$, la distribución complementaria acumulada ($P(R)$) y la integral de pérdida ($n(R)$) se expresan como sigue

$$P(R) = 0.5\exp(-[(R - \mu)/\theta])$$
$$n(R) = 0.5\theta\exp(-[(R - \mu)/\theta])$$

de modo que la relación $n(R)/P(R) = \theta$, independiente de R. Este hecho da como resultado una solución muy sencilla para el modelo (Q, R). Recuerde que las dos ecuaciones que definen la política óptima son

$$Q = \sqrt{\frac{2\lambda[K + pn(R)]}{h}}$$

$$P(R) = Qh/p\lambda \qquad [\text{donde } P(R) = 1 - F(R)].$$

La simplificación se logra usando la fórmula de la CPS presentada en la sección 5.5. Esta representación es una alternativa para Q que no incluye el costo p de faltantes. Usando $P(R) = 1 - F(R)$, la fórmula es

$$Q = \frac{n(R)}{P(R)} + \sqrt{\frac{2K\lambda}{h} + \left(\frac{n(R)}{P(R)}\right)^2}$$

$$= \theta + \sqrt{\frac{2K\lambda}{h} + \theta^2}$$

¡y es independiente de R! Por consiguiente, las Q y R óptimas se pueden calcular en una operación de una sola etapa. Cuando se usa un modelo de costo, se determina el valor de $P(R)$ a partir de $P(R) = Qh/p\lambda$. A continuación, usando la representación $P(R) = \exp[-(R - \mu)/\theta)]/2$ se llega a que $R = -\theta \ln(2P(R)) + \mu$. Si se utiliza un modelo de nivel de servicio, tan sólo se aplican las fórmulas de R dadas en la sección 5.5. Esto se muestra con un ejemplo.

Ejemplo *5D.3*

Veamos el ejemplo anterior 5D.2, pero supongamos que deseamos usar la distribución de Laplace para determinar la política óptima, y no la distribución de Poisson. Como la media y la varianza de la demanda en tiempo de demora son 2.8, igualamos $\mu = 2.8$ y $2\theta^2 = 2.8$, con lo que se obtiene $\theta = 1.1832$. Por consiguiente,

$$Q = 1.1832 + \sqrt{\frac{(2)(20)(72.8)}{10} + (1.1832)^2} = 18.3.$$

Como en el ejemplo anterior, obtenemos que $P(R) = .1934$. Usando $R = -\theta \ln(2P(R)) + \mu$ y al sustituir $P(R) = .1934$, se obtiene $R = 3.92$, que redondeamos a 4. Observe que esta solución es un poco distinta de la obtenida suponiendo la demanda con distribución de Poisson. Sin embargo, recuerde que al usar ésta, la R óptima estaba entre 4 y 5, y que elegimos el valor 5 para ser conservadores.

3. OTRAS DISTRIBUCIONES DE DEMANDA EN EL TIEMPO DE DEMORA

Se han recomendado muchas otras distribuciones de probabilidades para modelar la demanda en tiempo de demora. Entre otras están la distribución binomial negativa, la gamma, la logarítmica y la de Pearson. (Véase en Silver y Peterson, 1985, p. 289, una lista de artículos que describen esas distribuciones en el contexto de la administración de inventarios.) Es probable que la distribución normal sea la de mayor uso en las aplicaciones y que la de Poisson se implemente en casi todo el resto. Hemos incluido la distribución de Laplace por su interesante propiedad de que se pueden determinar las políticas óptimas (Q, R) sin una solución iterativa, y porque podría ser una buena alternativa cuando se trata de distribución de Poisson con artículos de poco movimiento.

Apéndice 5-E

Glosario de términos para el capítulo 5

α = Nivel deseado del servicio Tipo 1.

β = Nivel deseado del servicio Tipo 2.

c_o = Costo unitario de excedentes para el modelo del repartidor de periódicos.

c_u = Costo unitario de faltantes para el modelo del repartidor de periódicos.

D = Variable aleatoria que corresponde a la demanda. Es la demanda en un periodo para el modelo del repartidor de periódicos, y la demanda en el tiempo de demora para el modelo (Q, R).

CEP = Cantidad económica de pedido.

$F(t)$ = Función de distribución acumulada de la demanda. Los valores de esta función cuando la demanda tiene distribución normal figuran en la tabla A-4, al final del libro.

$f(t)$ = Función de densidad de probabilidades de la demanda.

$G(Q)$ = Costo esperado en un periodo asociado con el tamaño de lote Q (modelo del repartidor de periódicos).

$G(Q, R)$ = Costo anual promedio esperado para el modelo (Q, R).

h = Costo de (mantener el) inventario, por unidad de tiempo.

I = Tasa anual de interés empleada para calcular el costo de mantener el inventario.

K = Costo de preparación, o costo fijo de pedido.

μ = Tasa esperada de demanda (unidades por unidad de tiempo).

$L(z)$ = Función normalizada de pérdida. Se usa para calcular $n(R) = \sigma L(z)$. En la tabla A-4, al final del libro, aparecen valores tabulados de $L(z)$.

μ = Media de la demanda [demanda en el tiempo de demora para el modelo (Q, R)].

$n(R)$ = Cantidad esperada de faltantes en el tiempo de demora para el modelo (Q, R).

p = Costo de penalización por unidad, por no satisfacer la demanda.

Q = Tamaño (de lote) del pedido.

S = Existencia de seguridad, $S = R - \lambda\tau$ para el modelo (Q, R).

CPS = Cantidad de pedido para servicio.

T = Tiempo esperado del ciclo; tiempo promedio entre colocaciones de pedidos sucesivos.

τ = Tiempo de demora del pedido.

Servicio Tipo 1 = Proporción de ciclos en los que se satisface toda la demanda.

Servicio Tipo 2 = Proporción de las demandas satisfechas.

Bibliografía

AIN. "EDI Keeps Sportswear Shipments in Top Shape". *Automatic I.D. News*. Enero, 1995, pp. 38-39.

Arrow, K. A., T. E. Harris y T. Marschak. "Optimal Inventory Policy". *Econometrica* 19 (1951), pp. 250-272.

Arrow, K. A., S. Karlin y H. E. Scarf, eds. *Studies in the Mathematical Theory of Inventory and Production.* Stanford, CA: Stanford University Press, 1958.

Bessler, S. A. y A. F. Veinott, Jr. "Optimal Policy for a Dynamic Multiechelon Inventory Model". *Naval Research Logistics Quarterly* 13(1966), pp. 355-389.

Brown, R. G. *Statistical Forecasting for Inventory Control.* Nueva York: McGraw-Hill, 1959.

Brown, R. G. *Decision Rules for Inventory Management.* Hinsdale, IL: Dryden Press, 1967.

Clark, A. y H. E. Scarf. "Optimal Policies for a Multiechelon Inventory Problem". *Management Science* 6 (1960), pp. 475-490.

Cohen, M. A. y D. Pekelman. "LIFO Inventory Systems". *Management Science* 24 (1978), pp. 1150-1162.

Deuermeyer, B. L. y L. B. Schwarz. "A Model for the Analysis of System Service Level in Warehouse-Retailer Distribution Systems: Identical Retailer Case". En *Multilevel Production/Inventory Control Systems: Theory and Practice*, L. B. Schwarz, editor, pp. 163-194. Amsterdam: Holanda del Norte, 1981.

Dvoretsky, A., J. Kiefer y J. Wolfowitz. "The Inventory Problem: I. Case of Known Distributions of Demand". *Econometrica* 20(1952), pp. 187-222.

Ehrhardt, R. "The Power Approximation for Computing (s, S) Inventory Policies". *Management Science* 25 (1979), pp. 777-786.

Eppen, G. y L. Schrage. "Centralized Ordering Policies in a Multi-Warehouse System with Lead Times Random Demand". En *Multilevel Production/Inventory Systems: Theory and Practice*, L. B. Schwarz, editor, pp. 51-68. Amsterdam: Holanda del Norte, 1981.

Federgruen, A. y P. Zipkin. "Computational Issues in an Infinite Horizon Multi-Echelon Inventory Model". *Operations Research* 32 (1984), pp. 818-836.

Fishman, G. S. *Concepts and Methods in Discrete Event Digital Simulation.* Nueva York: John Wiley & Sons, 1973.

Freeland, J. R. y E. L. Porteus. "Evaluating the Effectiveness of a New Method of Computing Approximately Optimal (s, S) Inventory Policies". *Operations Research* 28 (1980), pp. 353-364.

Garry, M. "Completing the Loop". *Progressive Grocer* 74 (febrero 1995), pp. 75-80.

Graves, S. C. "A Multiechelon Inventory Model for a Repairable Item with One for One Replenishment". *Management Science* 31 (1985), pp. 1247-1256.

Graves, S. C., A. H. Rinnooy Kan G. y P. Zipkin, eds. *Handbooks in Operations Research and Management Science.* Volumen 4. *Logistics of Production and Inventory.* Amsterdam: Elsevier Science Publisher, 1992.

Hadley, G. J. y T. M. Whitin. *Analysis of Inventory Systems.* Englewood Cliffs, NJ: Prentice Hall, 1963.

Hartung, P. "A Simple Style Goods Inventory Model". *Management Science* 19 (1973), pp. 1452-1458.

Hausman, W. H. y R. Peterson. "Multiproduct Production Scheduling for Style Goods with Limited Capacity, Forecast Revisions, and Terminal Delivery". *Management Science* 18 (1972), pp. 370-383.

Iglehart, D. L. "Optimality of (s, S) Inventory Policies in the Infinite Horizon Dynamic Inventory Problem". *Management Science* 9 (1963), pp. 259-267.

Kaplan, R. "A Dynamic Inventory Model with Stochastic Lead Times". *Management Science* 16 (1970), pp. 491-507.

Love, S. F. *Inventory Control.* Nueva York: McGraw-Hill, 1979.

Mohan, S. "EDI's Move to Prime Time Stalled by Cost Perception". *Computerworld* 29 (20 de febrero de 1995), p. 91.

Muckstadt, J. M. y L. J. Thomas. "Are Multiechelon Inventory Models Worth Implementing in Systems with Low Demand Rate Items?" *Management Science* 26 (1980), pp. 483-494.

Murray, G. R. y E. A. Silver. "A Bayesian Analysis of the Style Goods Inventory Problem". *Management Science* 12 (1966), pp. 785-797.

Murray, J. E. "The EDI Explosion". *Purchasing* 118 (16 de febrero de 1995), pp. 28-30.

Nahmias, S. "Inventory Models". En *The Encyclopedia of Computer Science and Technology,* Volumen 9, J. Belzer, A. G. Holzman y A. Kent, editores, pp. 447-483. Nueva York: Marcel Dekker, 1978.

Nahmias, S. "Simple Approximations for a Variety of Dynamic Lead Time Lost-Sales Inventory Models". *Operations Research* 27 (1979), pp. 904-924.

Nahmias, S. "Managing Reparable Item Inventory Systems: A Review". En *Multilevel Production Inventory Control Systems: Theory and Practice*, L. B. Schwartz, editor, pp. 253-277. Amsterdam: Holanda del Norte, 1981.

Nahmias, S. "Perishable Inventory Theory: A Review". *Operations Research* 30 (1982), pp. 680-708.

Nahmias, S. y S. Smith. "Mathematical Models of Retailer Inventory Systems: A Review". En *Perspectives in Operations Management: Essays in Honor of Elwood S. Buffa*, R. K. Sarin, editor. Boston: Kluwer, 1992.

Parr, J. O. "Formula Approximations to Brown's Service Function". *Production and Inventory Management* 13 (1972), pp. 84-86.

Porteus, E. L. "Numerical Comparisons of lnventory Policies for Periodic Review Systems". *Operations Research* 33 (1985), pp. 134-152.

Presutti, V. y R. Trepp. "More Ado about EOQ". *Naval Research Logistics Quarterly* 17 (1970), pp. 243-251.

Ramberg, J. S. y B. W. Schmeiser. "An Approximate Method for Generating Symmetric Random Variables". *Communications of the ACM* 15 (1972), pp. 987-989.

Scarf, H. E. "The Optimality of (*s*, *S*) Policies in the Dynamic Inventory Problem". En *Mathematical Methods in the Social Sciences*, K. J. Arrow, S. Karlin y P. Suppes, editores. Stanford, CA: Stanford University Press, 1960.

Scarf, H. E. "Analytical Techniques in Inventory Theory". En *Multi-Stage Inventory Models and Techniques,* H. E. Scarf, D. M. Gilford y M. W. Shelly, editores. Stanford, CA: Stanford University Press, 1963.

Schmidt, C. P. y S. Nahmias. "Optimal Policy for a Single Stage Assembly System with Stochastic Demand". *Operations Research* 33 (1985), pp. 1130-1145.

Schwarz, L. B., ed. *Multilevel Production/Inventory Control Systems: Theory and Practice.* Amsterdam: Holanda del Norte, 1981.

Sherbrooke, C. C. "METRIC: Multiechelon Technique for Recoverable Item Control". *Operations Research* 16 (1968), pp. 122-141.

Silver, E. A. y R. Peterson. *Decision Systems for Inventory Management and Production Planning*, 2a. ed. Nueva York: John Wiley & Sons, 1985.

Veinott, A. F. "The Status of Mathematical Inventory Theory". *Management Science* 12 (1966), pp. 745-777.

Veinott, A. F. y H. M. Wagner. "Computing Optimal (*s*, *S*) Inventory Policies". *Management Science* 11 (1965), pp. 525-552.

Whitin, T. M. *The Theory of Inventory Management.* Ed. revisada. Princeton, NJ: Princeton University Press, 1957.

Capítulo **Seis**

Administración de la cadena de abastecimiento

Panorama general del capítulo

Propósito

Comprender qué son las modernas cadenas de abastecimiento, la forma en que se organizan y administran, y revisar los desarrollos más recientes en esta importante área.

Puntos clave

1. *¿Qué es una cadena de abastecimiento?* Una cadena de abastecimiento es toda la red relacionada con las actividades de una compañía que enlaza proveedores, fábricas, bodegas, almacenes y clientes. Requiere administración de artículos, dinero e información entre todos los participantes relevantes. El término *administración de la cadena de abastecimiento* (SCM, *supply chain management*) es relativamente nuevo, y evidentemente surgió de las experiencias de Procter and Gamble (P&G) a finales de los ochenta, cuando se rastreó el flujo de Pampers a través del canal de distribución. Sin embargo, manejar el flujo de artículos ya era un problema antes de acuñar la nomenclatura de administración de la cadena de abastecimiento. Habitualmente denominada logística, los problemas de la cadena de abastecimiento se han presentado desde la Revolución industrial. ¿Por qué este súbito interés en las cadenas de abastecimiento? Tradicionalmente, las compañías se enfocaban en la fabricación. Como resultado, la manufactura se ha vuelto relativamente eficiente, y hay pocas oportunidades de recortar costos de la función de fabricación. No obstante, existen todavía oportunidades significativas de reducir costos de la cadena de abastecimiento.
2. *El problema de transporte.* El problema de transporte es una de las primeras aplicaciones de la programación lineal. Supongamos que m se refiere a las instalaciones de producción (fuentes) y n a los puntos de la demanda (destinos). Se conoce el costo unitario de embarque de cada fuente a cada destino, y el objetivo es determinar un plan de embarque que satisfaga las restricciones de abastecimiento y demanda a un costo mínimo. La formulación de programación lineal del problema de transporte se ha resuelto exitosamente con cientos de miles de variables y restricciones. Una generalización del problema de transporte es el problema de trasbordo. Aquí pueden utilizarse nodos intermedios para el almacenaje, así como puntos de abastecimiento o demanda. Los problemas de trasbordo se pueden solucionar también con programación lineal.
3. *Planeación de recursos de distribución y envío.* En la sección 6.5 proporcionamos una breve discusión sobre la planeación de recursos de distribución (DRP, *distribution resource planning*). DRP es la aplicación de la planeación de requerimientos de materiales (MRP, comentada en el capítulo 7) a los problemas de distribución. El método tiene ventaja

sobre el enfoque del punto de reorden (ROP) comentado en el capítulo 5 en que permite una variación predecible en el patrón de demanda. Sin embargo, no toma en cuenta la variación no sistemática (aleatoriedad) estudiada de manera explícita en el análisis ROP.

La sección 6.6 ofrece una breve introducción a los problemas en la ruta de vehículos. Consideremos un camión de reparto que debe hacer entregas a varios clientes. El objetivo es encontrar la secuencia óptima de entregas que minimice la distancia total requerida. Esto se conoce como el problema del agente viajero, y resulta muy difícil de resolver de forma óptima. Los cálculos que se requieren para encontrar la solución óptima crecen exponencialmente con el tamaño de problema (conocido en matemáticas como problema difícil NP). En esta sección presentamos una heurística sencilla para obtener soluciones aproximadas, conocida como método de ahorros.

4. *Diseño de productos para la eficiencia de la cadena de abastecimiento.* El "enfoque innovador" se ha convertido en un cliché. Significa ver el problema de una nueva forma, con frecuencia sin tomar las restricciones tal como están planteadas. Un ejemplo de enfoque innovador es la posposición en las cadenas de abastecimiento. La primera aplicación de esta idea se debe a Benetton, el conocido fabricante de ropa tejida. Benetton debe predecir las preferencias de color de los consumidores con anticipación a la temporada de venta. Debido a que la lana se tiñe primero y después se teje para convertirla en suéteres, la mezcla de color se debe decidir con mucha anticipación. Si sus predicciones sobre las referencias de color de los consumidores están equivocadas (lo que siempre sucede), los colores populares se venderán con mucha rapidez y los que no lo son se quedarán en los anaqueles. Su solución fue revertir el orden de las operaciones de tejido y teñido. Los suéteres se tejieron con lana sin teñir (gris) y después se teñían lo más tarde posible. Esto proporcionó más tiempo para observar qué colores se vendían mejor. Hewlett-Packard descubrió una solución similar en su división de impresoras. Éstas deben configurarse para los mercados locales debido al idioma y otras diferencias. Al producir impresoras "grises" con todas las partes en común, y después configurar las de exportación *in situ* en los mercados locales, pudieron retrasar la diferenciación del producto y balancear mejor sus inventarios.

   Otro ejemplo de diseño de productos para la eficiencia de la cadena de abastecimiento es Ikea. Ikea es la compañía sueca de muebles económicos para el hogar. Para reducir costos, diseña su mobiliario para poder almacenarlo con facilidad directamente en los establecimientos comerciales. Esto significa que los clientes se pueden llevar su compra consigo, evitando así las largas esperas y personalización requerida por las mueblerías más tradicionales.

5. *El papel de la información en las cadenas de abastecimiento.* Como se comentó, una cadena de abastecimiento involucra la transferencia de artículos, dinero e información. La administración moderna de la cadena de abastecimiento busca eliminar ineficiencias surgidas de los flujos de información deficientes. Una forma de reducir este problema es por medio de inventarios administrados por el proveedor. Tanto Wal-Mart como Barilla (un productor italiano de pasta) se han beneficiado de ellos. En estas situaciones, los proveedores, más que los vendedores, son responsables de mantener el inventario en los anaqueles.

   Como se comentó antes, este término surgió de la experiencia de seguimiento de P&G acerca de la variabilidad de las ventas de Pampers a través de las diferentes etapas de la cadena de abastecimiento. Lo que P&G observó fue que aun cuando la demanda del consumidor del producto final se mantuvo prácticamente igual con el tiempo, había grandes variaciones en el patrón de pedidos a la fábrica. Varios han especulado acerca de las causas de este fenómeno. Una consecuencia similar aparece cuando se juega el Beer Game, una simulación concebida por Ray Forrester.

   El comercio electrónico y el intercambio electrónico de datos (IED) han avanzado mucho para hacer más eficientes las operaciones de la cadena de abastecimiento al

mejorar la velocidad y eficiencia de las transacciones. El IED incluye sistemas de intercambio de datos privados y sistemas de transacciones basadas en Internet.

6. *Sistemas de distribución multinivel.* En los grandes sistemas es común almacenar los artículos en varios lugares. Los centros de distribución (CD) reciben productos de las plantas y fábricas y luego los envían a centros más pequeños o directamente a las tiendas. Algunas de las ventajas de emplear CD incluyen las economías de escala, ajustes en la mezcla de producto a una región o cultura en particular y reducir el inventario de seguridad por medio de la cooperativa de riesgo.
7. *Diseño de la cadena de abastecimiento en un ambiente global.* Hoy en día, la mayor parte de las compañías son multinacionales. Los productos se diseñan y embarcan para una amplia variedad de mercados en todo el mundo. Como ejemplo consideremos el mercado de automóviles. Hace cincuenta años, casi todos los que se vendían en Estados Unidos se produjeron en ese país. Actualmente el porcentaje se ha reducido a cerca del 50 por ciento. Las fuerzas del mercado global están forjando la nueva economía. Surgen vastos mercados, como China, y los principales poderes industriales están deseosos de obtener su parte. La tecnología, las consideraciones de costos y las fuerzas políticas y macroeconómicas han dirigido la globalización. La venta en diversos mercados genera problemas especiales en la administración de la cadena de abastecimiento.

Vamos a la tienda de abarrotes a las 10:00 p.m. a comprar un bote de crema de cacahuate para el almuerzo escolar de nuestro hijo al día siguiente. No sólo está abierta sino que hay una gran variedad, marcas, estilos y tamaños disponibles de crema de cacahuate. Los estadounidenses (y los residentes de los países más modernos) dan esas cosas por sentadas. Sin embargo, existe una gran cantidad de complejas actividades que deben coordinarse cuidadosamente para asegurar que la crema de cacahuate esté ahí cuando la necesitamos. Y es lo mismo en el caso de la ropa, equipo de cómputo y todos los otros bienes de consumo que compramos. ¿Valoramos el que muchos de ellos se producen y embarcan desde todo el mundo antes de llegar a nuestros hogares? La logística para coordinar las actividades que nos proporcionan esta conveniencia es la esencia de la administración de la cadena de abastecimiento.

El término *administración de la cadena de abastecimiento* (SCM) parece haber surgido a finales de los ochenta y sigue captando el interés en un porcentaje muy elevado. La literatura comercial está llena de libros y artículos que se relacionan con alguno de sus aspectos. Hoy en día son muy comunes el software y las firmas consultoras especializadas en SCM. Éstas han crecido de manera muy importante e incluyen a los gigantes SAP y Oracle, que ofrecen soluciones SCM como parte de sistemas completos de recuperación de información. Aunque la etiqueta SCM es relativamente nueva, los problemas considerados no lo son. Casi todo el material de los capítulos 2 a 5 se relaciona con la SCM. Así, ¿qué es lo novedoso de SCM? La sencilla respuesta es que ésta ve el problema del manejo de artículos como un sistema integrado. Se ha propuesto una gran cantidad de definiciones para la SCM, y resulta aleccionador examinar algunas de ellas. La más sencilla y directa aparece en la página de Internet del Foro de la Cadena de Abastecimiento de Stanford (1999) y se debe probablemente a Hau Lee, su director. Dice:

> La administración de la cadena de abastecimiento tiene que ver con el manejo de materiales, información y flujos financieros en una red consistente de proveedores, fabricantes, distribuidores y clientes.

Aunque breve, esta definición es prácticamente completa. Indica que lo importante no es sólo el flujo de productos sino también el de información y dinero.

Una definición realizada por Simchi-Levi *et al*. (1999, p. 1) que se enfoca únicamente en el flujo de productos es:

> La administración de la cadena de abastecimiento es un conjunto de enfoques utilizado para integrar a los proveedores, fabricantes, bodegas y almacenes de manera eficiente, para que la mercancía se produzca y distribuya en las cantidades adecuadas, a los lugares correctos, y de manera oportuna, a fin de minimizar los costos de todo el sistema al tiempo que se cubren los requerimientos de nivel de servicio.

Como esta definición se enfoca en el flujo de productos, implica que la SCM sólo es relevante para las compañías fabricantes. ¿Lo es también para las organizaciones de servicio? ¿Dónde intervienen las cuestiones de estrategia en estas definiciones? A las claras es difícil emitir una definición que integre todos los elementos. Sin embargo, formularla de forma general en exceso oscurece el significado. Consideremos, por ejemplo, la siguiente definición realizada por Ross (1998):

> La administración de la cadena de abastecimiento es una filosofía de administración que evoluciona de manera continua, busca unificar las competencias productivas colectivas y los recursos de las funciones empresariales dentro de la empresa y fuera de ella con los socios comerciales aliados, localizados a lo largo de canales de abastecimiento intersecados dentro de un sistema de abastecimiento enriquecedor para el cliente y altamente competitivo enfocado en desarrollar soluciones y sincronizar el flujo de productos de mercado, servicios e información para crear fuentes individualizadas de valor para el cliente.

Quizás la descripción más clara sea la que apareció en un artículo de la revista *Fortune* dedicado al tema:

> Llámela distribución o logística o administración de la cadena de abastecimiento. Cualquiera que sea el nombre, es el sinuoso, molesto y denodado proceso por el que deben pasar las compañías para llevar material, partes y productos a los clientes. En industria tras industria, desde autos y ropa hasta computadoras y productos químicos, los ejecutivos han arrancado ésta, alguna vez una deprimente disciplina de la zona de carga, llevándola cerca de la parte más alta de la agenda corporativa. Presionadas para acabar con sus competidores en calidad o precio, las compañías tratan de obtener una ventaja a través de su capacidad para proporcionar el material adecuado en la cantidad apropiada y en el momento oportuno. (Henkoff, 1994)

La mayor parte de los escritores acuerdan que la "logística" se relaciona en esencia con las mismas cuestiones que la administración de la cadena de abastecimiento. El término *logística* se originó en los problemas militares de mover personal y material en épocas críticas. Las empresas lo adoptaron más tarde y se convirtió en un apelativo común en las sociedades profesionales y programas académicos. Podría pensarse que aunque la logística clásica trató el mismo conjunto de problemas que la SCM, no consideró a la cadena de abastecimiento como un sistema integrado (Copacino, 1997). En la figura 6-1 aparece un resumen de la "sombrilla" de actividades que componen la SCM.

En un sentido, podemos considerar que todo este texto tiene que ver con algún aspecto de SCM si tomamos el panorama completo del campo que aparece en la figura 6-1. Sin embargo,

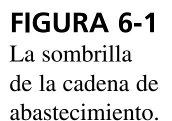
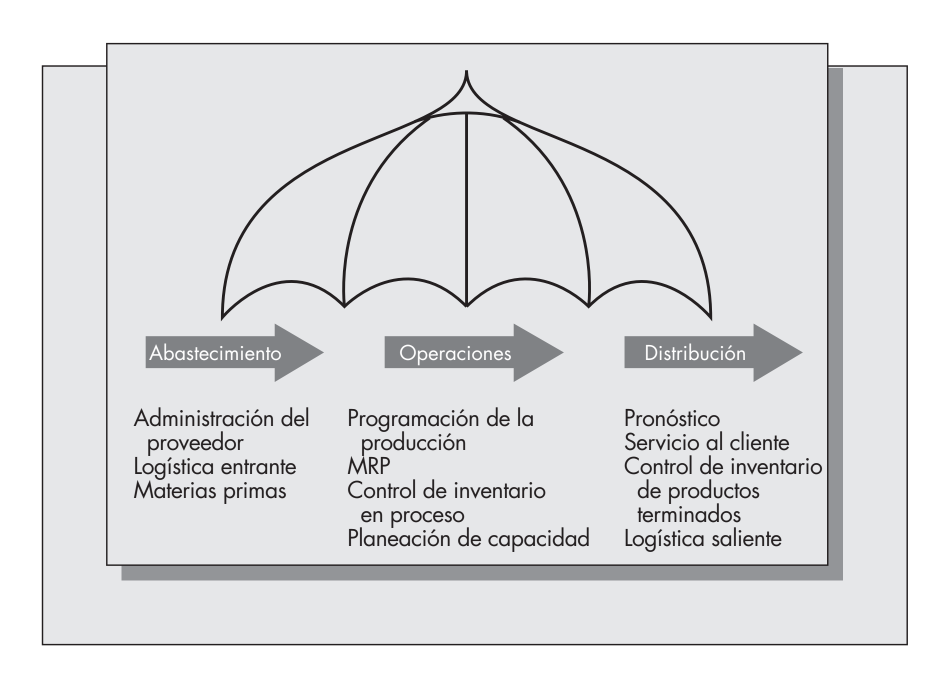


FIGURA 6-1
La sombrilla de la cadena de abastecimiento.



la mayoría de los capítulos se enfocan en temas específicos, como el control de inventarios o la programación de trabajos en el taller. La SCM, por otro lado, se refiere a *todo* el sistema de entrega del producto desde los proveedores de materias primas hasta los canales de distribución de los artículos terminados. Aunque importantes y útiles en varios contextos, es improbable que fórmulas sencillas como la CEP arrojen mucha luz sobre la administración efectiva de cadenas de abastecimiento complejas.

La cadena de abastecimiento como un arma estratégica

¿Dónde encaja la cadena de abastecimiento en la estrategia global del negocio? En el capítulo 1 se comentó que parte importante de la ventaja competitiva radicaba en su posicionamiento estratégico en el mercado. Algunos ejemplos de éste son ser el proveedor de más bajo costo (como los automóviles Hyundai) o ser proveedor de alta calidad (como Mercedes-Benz) o explotar la segmentación de mercado para obtener ganancias en ambos mercados (como lo hace General Motors con diferentes marcas dirigidas a distintos segmentos de mercado). El diseño de la cadena de abastecimiento refleja también el posicionamiento estratégico de la compañía.

En una cadena de abastecimiento, la principal interrelación se encuentra entre el costo y el tiempo de respuesta. Es menos costoso transportar el producto por barco o camión, pero el transporte aéreo puede llevar el producto del punto A al punto B con mucha mayor rapidez. ¿Serán más confiables las entregas si el producto se mueve por medio del sistema interno de la compañía, o será mejor subcontratar la operación de logística a un tercero? La logística de terceros (que se abrevia como 3PL) está volviéndose cada vez más común. ¿Por qué? Por la misma razón de la gran expansión de la fabricación por terceros. Las compañías como Solectron Corporation, de Silicon Valley, pueden lograr significativas economías de escala brindando servicios de fabricación a grandes cantidades de compañías que elaboran productos similares. De esta forma, subcontratar la fabricación quizás cueste menos que realizarla de manera interna. El costo no es el único problema. Por ejemplo, Trimble Corporation, productores de sistemas de posicionamiento global de vanguardia, ha debido enfrentar que el diseño del producto, no la fabricación, era la competencia central de la compañía. A finales de los noventa Trimble cerró sus instalaciones de fabricación y hoy en día subcontrata la mayor parte de la fabricación con Solectron. (En una posible coincidencia, el precio de las acciones de Trimble se elevó de forma muy importante después de que la compañía realizó ese cambio.)

La función de fabricación se ha vuelto muy eficiente en los últimos años. Las compañías han optado por establecerla como una competencia clave, o bien la subcontratan. Por esa razón, la fabricación por lo común no proporciona una oportunidad para reducir los costos de modo importante. ¿De dónde, entonces, provienen las reducciones de costo si no es de ella? La respuesta es de la cadena de abastecimiento. Las compañías que son capaces de mover el producto con rapidez y eficiencia obtendrán una ventaja sobre sus competidores. Nuevas tecnologías, como los sistemas de administración de abastecimiento con base en Internet (que se comentan más adelante en la sección 6.8), y la aplicación inteligente del modelado matemático, son las formas en que las compañías reducirán costos de la cadena de abastecimiento sin comprometer el servicio.

La Aplicación instantánea que se presenta en esta sección habla sobre Wal-Mart, cuyo extraordinario éxito se logró hasta cierto punto debido a sus sofisticadas estrategias en la cadena de abastecimiento.

Hay dos metas en este capítulo. La primera es proporcionar al lector una apreciación de los problemas más importantes encontrados al manejar cadenas de abastecimiento complejas. El segundo es presentar una muestra de los tipos de modelos matemáticos usados en el análisis de SCM.

6.1 EL PROBLEMA DE TRANSPORTE

El problema de transporte es un modelo matemático para calendarizar de forma óptima el flujo de productos de las instalaciones de fabricación a los centros de distribución. Supongamos que una cantidad fija de producto debe transportarse de un grupo de fuentes (plantas) a un gru-

Aplicación instantánea

WAL-MART GANA CON UNA SÓLIDA ADMINISTRACIÓN DE SU CADENA DE ABASTECIMIENTO

Aunque existen muchos ejemplos de compañías que ganan (o pierden) por sus buenas o malas estrategias de cadena de abastecimiento, quizás ninguno es más espectacular que las historias de Wal-Mart y Kmart. Ambas compañías se fundaron en el mismo año, 1962. En unos años Kmart se convirtió en un nombre muy familiar, mientras Wal-Mart era prácticamente desconocido a excepción de algunas comunidades en el sur. Las tiendas Wal-Mart se ubicaron normalmente en áreas rurales, así que rara vez competían con las cadenas más grandes y mejor conocidas como Sears y Kmart, cuyas tiendas en su mayoría estaban ubicadas en grandes ciudades y en los suburbios cercanos. En 1987 las ventas de Kmart eran casi el doble de las de Wal-Mart (cerca de $25 millardos anuales contra $15 millardos anuales). Para 1990 Wal-Mart había alcanzado a Kmart y en 1994 sus ventas anuales casi triplicaron las de Kmart (unos $80 millardos contra $27 millardos). Wal-Mart es ahora el minorista de descuento más grande de Estados Unidos, sobrepasando a Sears y Kmart. ¿Cuál puede haber sido la causa para que se registrara este giro tan impresionante?

Aunque algunos pueden señalar varios factores que llevaron al éxito de Wal-Mart, no hay duda de que el énfasis de la compañía en una sólida administración de la cadena de abastecimiento fue uno de los más importantes. De acuerdo con Duff y Ortega (1995), al comparar las decisiones del antiguo CEO de Kmart, Joseph Antonini, y del de Wal-Mart, Sam Walton:

> Cuando el señor Antonini tomó las riendas de Kmart en 1987, tenía las manos llenas. Además, sus antecesores se rehusaron a implementar los sofisticados sistemas computacionales que ayudaban a Wal-Mart a dar seguimiento y reabastecer la mercancía de manera rápida y eficiente.
>
> Un autopromotor de voz fuerte y amplia sonrisa, el señor Antonini, invirtió pródigamente en campañas de televisión a escala nacional y glamorosos representantes como Jaclyn Smith. El señor Walton evitó la publicidad. En lugar de enfocarse en la mercadotecnia, se obsesionó con las operaciones. Invirtió decenas de millones de dólares en un sistema de computadora para toda la compañía que unía las cajas registradoras con las oficinas generales, permitiéndole reabastecer rápidamente en los anaqueles todos los artículos que se vendían. Invirtió también una gran cantidad en camiones y centros de distribución. Además de mejorar su control, esto redujo bastante los costos.
>
> El señor Antonini trató de estimular el crecimiento supervisando la compra de otros tipos de minoristas: los artículos deportivos Sports Authority, las tiendas de artículos de oficina OfficeMax, las librerías Borders y los clubes Pace Membership Warehouse.
>
> Sin embargo, la diferencia menos visible entre Wal-Mart y Kmart empezaba a ser de gran importancia. Los increíblemente sofisticados sistemas de lectura y distribución de inventario dieron por resultado que los clientes nunca encontraran los anaqueles vacíos o retrasos en la revisión de precios de los artículos en las cajas registradoras.
>
> Los pasillos de Kmart, mientras tanto, estaban llenos de historias de terror en cuanto a la distribución. Joseph R. Thomas, quien la supervisó, comentó que, en retrospectiva, debía haberse percatado de que había problemas cuando encontró las bodegas llenas de mercancía el 15 de diciembre, el apogeo de la temporada navideña.

Aquí tenemos importantes lecciones. Éstas van al centro de lo que ha estado equivocado en el pensamiento comercial estadounidense durante décadas. Cuando los problemas surgieron, ¿cómo respondió la administración de Kmart? Gastando dinero en mercadotecnia y adquisiciones. Wal-Mart, por otro lado, invirtió en mejores tiendas y sistemas de logística que le aseguraran que los estantes de dichas tiendas estuvieran abastecidos con lo que los clientes deseaban. Hoy Wal-Mart tiene uno de los sistemas de intercambio electrónico de datos (IED) y control de inventarios más sofisticados de la industria.

La historia de Kmart y Wal-Mart recuerda lo ocurrido con los autos japoneses y estadounidenses. En los cincuenta y sesenta, cuando los fabricantes de automóviles estadounidenses tenían un virtual monopolio en gran parte del mercado de automóviles del mundo, ¿dónde colocaron sus recursos los Tres Grandes? ¿En producir autos mejor diseñados a un menor costo? No, en campañas llamativas y remodelaciones cosméticas anuales. Por su parte, los japoneses invirtieron en lo más reciente de la tecnología automotriz y en sofisticados sistemas de logística como el justo-a-tiempo. El resto, como dicen, es historia.

po de depósitos (bodegas). Asumimos conocer el costo unitario de transportar de cada fuente a cada depósito. La meta es encontrar las rutas óptimas de flujo y las cantidades que se van a enviar en esas rutas para minimizar el costo total de todos los embarques.

El problema de transporte puede verse como un problema prototipo de la cadena de abastecimiento. Aunque la mayoría de los problemas del mundo real relacionados con el embarque de los productos tiene mayor complejidad, el modelo proporciona una imagen de los problemas y métodos que podemos encontrar en la práctica.

Ejemplo *6.1*

La Pear Disk Drive Corporation produce varias capacidades de unidades de disco para computadoras personales. En 1999, las fabricó con capacidades de 20 a 160 gigabytes (GB), todas en el factor de forma de 3.5 pulgadas. El producto más popular es la unidad de 80 GB, que se vende a varios fabricantes de computadoras. Pear las produce en tres plantas ubicadas en Sunnyvale, California; Dublín, Irlanda, y Bangkok, Tailandia. Periódicamente se realizan los embarques desde estos tres centros de producción a cuatro bodegas de distribución ubicadas en Estados Unidos en Amarillo, Texas; Teaneck, Nueva Jersey; Chicago, Illinois, y Sioux Falls, Dakota del Sur. Durante el mes próximo, se ha determinado que estas bodegas deben recibir las siguientes proporciones de la producción total de unidades de disco de 80 GB de la compañía:

Bodega	Porcentaje de producción total
Amarillo	31
Teaneck	30
Chicago	18
Sioux Falls	21

Se espera que las cantidades de producción en las fábricas durante el siguiente mes sean (en miles de unidades)

Planta	Producción anticipada (en miles de unidades)
Sunnyvale	45
Dublín	120
Bangkok	95

Como el total de la producción en las tres plantas es de 260 unidades, la cantidad embarcada a las cuatro bodegas será (redondeando a la unidad más cercana).

Bodega	Cantidad total del embarque (miles)
Amarillo	80
Teaneck	78
Chicago	47
Sioux Falls	55

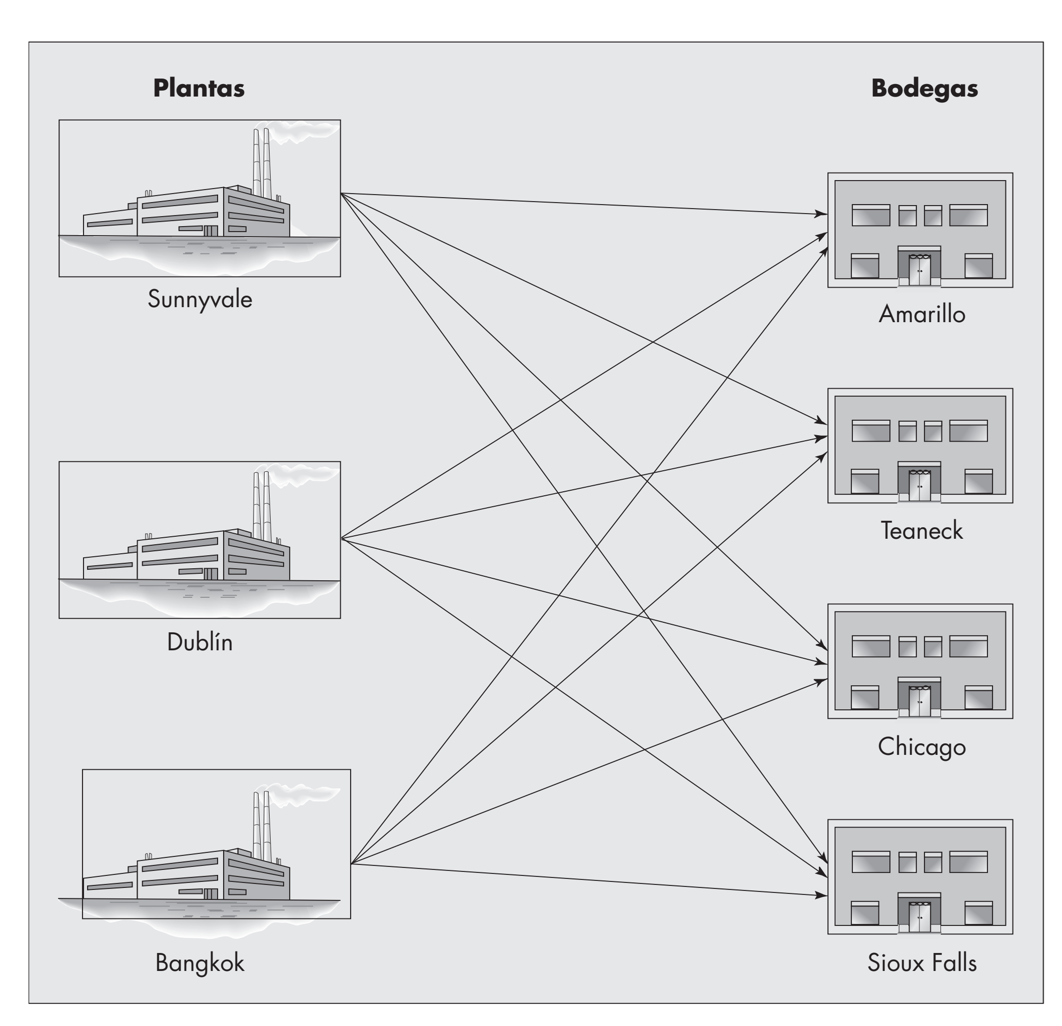
A pesar de que el costo de embarque puede ser menor entre ciertas plantas y los centros de distribución, Pear ha establecido rutas de embarque entre cada planta y cada bodega. Esto en caso de existir problemas imprevistos como el cierre forzado en una planta, vaivenes no anticipados en las demandas

regionales, o mal tiempo entre algunas rutas. Los costos unitarios de embarcar 1 000 unidades de cada planta a cada bodega se muestran en la siguiente tabla:

Costos de embarque por mil unidades en $					
		A			
		Amarillo	**Teaneck**	**Chicago**	**Sioux Falls**
DE	**Sunnyvale**	250	420	380	280
	Dublín	1 280	990	1 440	1 520
	Bangkok	1 550	1 420	1 660	1 730

La meta es determinar un patrón de embarque que minimice el costo total de transportación de las plantas a las bodegas. La representación de la red del problema de distribución de Pear se muestra en la figura 6-2.

FIGURA 6-2
Problema de transportación de Pear Disk Drive.



La heurística avara

El problema de transportación puede formularse como un programa lineal y por tanto resolverse al utilizar un código de programación lineal, como Solver en Excel. Para adquirir cierta intuición acerca de la estructura del problema, sin embargo, consideraremos una heurística sencilla que por lo general da buenas soluciones, aunque posiblemente por debajo de lo óptimo. Para implementarla, construimos una tabla de transporte.

Ejemplo *6.1* *(continuación)*

La tabla de transporte es parecida a la primera tabla que proporciona los costos de embarque de las plantas a las bodegas. Esto es

	Bodegas				
Fábricas	Amarillo	Teaneck	Chicago	Sioux Falls	
Sunnyvale	250	420	380	280	45
Dublín	1 280	990	1 440	1,520	120
Bangkok	1 550	1 420	1 660	1 730	95
	80	78	47	55	

Observe la estructura de la tabla. Las filas corresponden a las fuentes de abastecimiento (fábricas) y las columnas a los depósitos (bodegas). Los números que colocaremos en las celdas del cuadro serán el valor del flujo de producto de cada fuente a cada depósito. Para implementar la heurística avara, buscamos el cuadro para el costo unitario mínimo, que es $250 y corresponde a la celda Sunnyvale/Amarillo. En esta celda colocamos el mínimo de las siguientes cantidades: la disponibilidad en Sunnyvale (45) y el requerimiento en Amarillo (80). Por tanto, colocamos 45 en esta celda. En este punto, hemos saturado la primera fila (quiere decir que no puede embarcarse más desde Sunnyvale), así que dibujamos una línea a lo largo de la primera fila. La tabla se convierte ahora en

	Bodegas				
Fábricas	Amarillo	Teaneck	Chicago	Sioux Falls	
Sunnyvale	250 ~~45~~	420	380	280	~~45~~
Dublín	1280	990	1440	1520	120
Bangkok	1550	1420	1660	1730	95
	80	78	47	55	

De las celdas descubiertas, la de menor costo es Dublín/Teaneck en $990 [es decir, la celda (2,2)]. Aquí podemos asignar un flujo máximo de 78 unidades (la mínima entre 78 y 120), lo que hacemos. Esto satura la segunda columna, así que trazamos una línea a lo largo de la segunda columna para indicarlo. La siguiente celda de menor costo es Dublín/Amarillo a $1 280. La capacidad de esta celda es

mínima (80 − 45, 120 − 78), ya que 45 está asignado a la primera columna y 78 a la segunda. Por tanto, asignamos mín(35, 42) = 35 a esta celda, que satura la primera columna. La tabla en esta etapa es

Fábricas	**Bodegas** Amarillo	Teaneck	Chicago	Sioux Falls	
Sunnyvale	250 / ~~45~~	420	380	280	~~45~~
Dublín	1 280 / 35	990 / 78	1 440	1 520	120
Bangkok	1 550	1 420	1 660	1 730	95
	~~80~~	78	47	55	

La heurística avara continúa de esta manera hasta que estén saturadas todas las columnas. La solución final es

Fábricas	**Bodegas** Amarillo	Teaneck	Chicago	Sioux Falls	
Sunnyvale	250 / ~~45~~	420	380	280	~~45~~
Dublín	1 280 / 35	990 / 78	1 440 / 7	1 520	120
Bangkok	1 550	1 420	1 660 / 40	1 730 / 55	95
	80	78	47	55	

Si aceptamos que x_{ij} sea la cantidad de flujo desde la fuente i hasta el depósito j, entonces la solución que se muestra en la última tabla es

$$x_{11} = 45, \quad x_{21} = 35, \quad x_{22} = 78, \quad x_{23} = 7,$$
$$x_{33} = 40, \quad \text{y} \quad x_{34} = 55,$$

y todos los otros $x_{ij} = 0$. El costo total de esta solución es (45)(250) + (35)(1 280) + · · · = $304 900. Como veremos en la sección 6.2, esta solución es menos que óptima para este problema.

6.2 RESOLUCIÓN DE PROBLEMAS DE TRANSPORTE CON PROGRAMACIÓN LINEAL

Se han propuesto varias heurísticas para resolver los problemas de transporte, como el que se presenta en la sección 6.1. No obstante, es poco probable que alguien con un problema real pueda usar una heurística, ya que las soluciones óptimas se pueden encontrar eficientemente

por medio de la programación lineal. De hecho, a causa de la estructura particular del problema de transporte, los códigos especializados de hoy en día pueden resolver los problemas con millones de variables. La heurística avara se presentó en la sección 6.1 para dar al lector una idea de la forma en que se asignan los flujos en las redes y la forma en que se pueden obtener ágilmente soluciones razonables.

Sea m es el número de fuentes y n el número de depósitos. (En el ejemplo 6.1, $m = 3$ y $n = 4$.) Recordemos la definición de las variables de decisión de la sección 6.1:

$$x_{ij} = \text{flujo desde la fuente } i \text{ hasta el depósito } j \text{ para } 1 \le i \le m \text{ y } 1 \le j \le n,$$

y definamos c_{ij} como el costo de embarcar una unidad de i a j. Se deduce que el costo total de realizar todos los embarques es

$$\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

Para el caso de Pear Company, descrito en el ejemplo 6.1, la función objetivo es

$$250x_{11} + 420x_{12} + 380x_{13} + 280x_{14} + \cdots + 1730x_{34}.$$

Como es obvio que muchas rutas no son económicas, es probable que muchas de las variables de decisión sean iguales a cero en la solución óptima.

Las restricciones están diseñadas para asegurar que la cantidad total embarcada de cada fuente sea igual a la cantidad allí disponible, y la cantidad embarcada hacia cada depósito sea igual a la cantidad requerida por ese depósito. Dado que hay m fuentes y n depósitos, existe un total de $m + n$ restricciones (excluyendo las restricciones de no negatividad). Digamos que a_i es la cantidad total que se va a embarcar desde la fuente i, y b_j la cantidad total que se va a embarcar al depósito j. Las restricciones de programación lineal se pueden escribir:

$$\sum_{j=1}^{n} x_{ij} = a_i \qquad \text{para } 1 \le i \le m$$

$$\sum_{i=1}^{m} x_{ij} = b_j \qquad \text{para } 1 \le j \le n$$

$$x_{ij} \ge 0 \qquad \text{para } 1 \le i \le m \quad \text{y} \quad 1 \le j \le n.$$

Para el problema de Pear Disk Drive Company, obtenemos las siguientes siete restricciones:

$$\begin{aligned}
x_{11} + x_{12} + x_{13} + x_{14} &= 45 && \text{(embarques desde Sunnyvale)}\\
x_{21} + x_{22} + x_{23} + x_{24} &= 120 && \text{(embarques desde Dublín)}\\
x_{31} + x_{32} + x_{33} + x_{34} &= 95 && \text{(embarques desde Bangkok)}\\
x_{11} + x_{21} + x_{31} &= 80 && \text{(embarques hacia Amarillo)}\\
x_{12} + x_{22} + x_{32} &= 78 && \text{(embarques hacia Teaneck)}\\
x_{13} + x_{23} + x_{33} &= 47 && \text{(embarques hacia Chicago)}\\
x_{14} + x_{24} + x_{34} &= 55 && \text{(embarques hacia Sioux Falls)}
\end{aligned}$$

y las restricciones de no negatividad requeridas en la programación lineal:

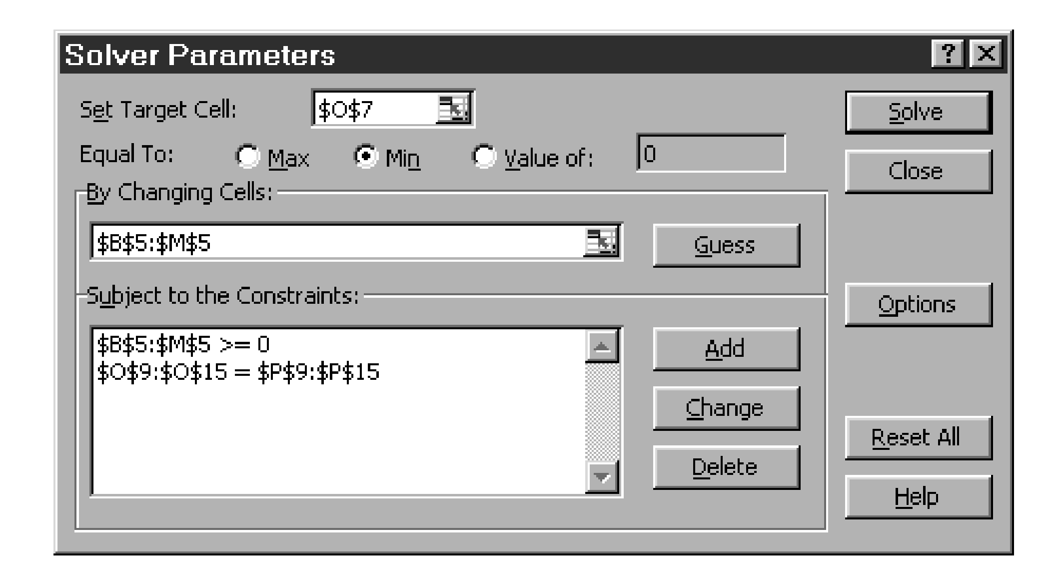
$$x_{ij} \ge 0 \qquad \text{para } 1 \le i \le 3 \quad \text{y} \quad 1 \le j \le 4.$$

El problema se introdujo en el Solver de Excel. La hoja de cálculo utilizada y la solución aparecen en la figura 6-3. La solución obtenida es

$$x_{14} = 45, \qquad x_{21} = 42, \qquad x_{22} = 78, \qquad x_{31} = 38,$$
$$x_{33} = 47 \quad \text{y} \quad x_{34} = 10,$$

FIGURA 6-3
Solución al problema de transportación de Pear utilizando Solver de Excel.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1			**Solución del ejemplo 6.1 (Problema Pear's Transportation)**													
2																
3	Variables	×11	×12	×13	×14	×21	×22	×23	×24	×31	×32	×33	×34	Operador	Valor	RHS
4																
5	Valor	0	0	0	45	42	78	0	0	38	0	47	10			
6																
7	Coef. objetivo	250	420	380	280	1280	990	1440	1520	1550	1420	1660	1730	Min	297800	
8	st															
9	Restricción 1	1	1	1	1									=	45	45
10	Restricción 2					1	1	1	1					=	120	120
11	Restricción 3									1	1	1	1	=	95	95
12	Restricción 4	1				1				1				=	80	80
13	Restricción 5		1				1				1			=	78	78
14	Restricción 6			1				1				1		=	47	47
15	Restricción 7				1				1				1	=	55	55
16																
17																
18		Notas:	Fórmula para la celda O9: =SUMPRODUCT (B9:M9,$B$5:$M$5). Copiado de O10 a O15.													
19			Las celdas del Solver que cambian son $B$5:$M$5.													
20																



con todos los otros valores iguales a cero. El costo total de esta solución es $297 800. Observe que no es igual a la solución obtenida de la heurística avara. Como se observó antes, el costo de la heurística avara es de $304 900, lo que difiere del costo óptimo en sólo 2.4 por ciento.

6.3 ASPECTOS GENERALES DEL PROBLEMA DE TRANSPORTE

El ejemplo de Pear Company es el tipo más sencillo de problema de transporte. Es factible cada enlace desde una fuente a un depósito, y la cantidad total disponible de las fuentes es exactamente igual a la demanda total en los depósitos. Varios de estos requerimientos pueden relajarse sin necesidad de complicar el problema significativamente.

Rutas no factibles

Supongamos en el ejemplo que las compañías han decidido eliminar las rutas de Dublín a Chicago y de Bangkok a Sioux Falls. Esto se justificará colocando costos muy altos en estos arcos de la red. Tradicionalmente, la *A* mayúscula se ha utilizado para representar un costo muy elevado. En la práctica, desde luego, debe asignarse un número a estas ubicaciones. Debido a que éste es mucho mayor que el resto de los costos, una solución óptima nunca asignará un flujo a estas rutas. Por ejemplo, supongamos que asignamos costos de $1 000 000 a cada una de estas rutas y resolvemos el problema. El lector puede verificar que ahora obtenemos la siguiente solución:

$$x_{14} = 45, \quad x_{21} = 32, \quad x_{22} = 78, \quad x_{31} = 48 \quad \text{y} \quad x_{33} = 47,$$

con todos los otros valores iguales a cero. El costo de la nueva solución es $298 400, apenas ligeramente mayor que el costo obtenido cuando todas las rutas eran factibles.

Problemas desbalanceados

Un problema de transporte desbalanceado es aquel donde la cantidad total embarcada desde las fuentes no es igual a la cantidad total requerida en los depósitos. Esto puede surgir si la demanda excede el abastecimiento o viceversa. Hay dos formas de manejar los problemas desbalanceados. Una consiste en sumar una fila o columna ficticias para absorber la demanda o el abastecimiento excesivos. Un segundo método para resolver los problemas desbalanceados es alterar el conjunto adecuado de restricciones para cualquier forma $\leq$ o $\geq$. Ambos métodos se ilustrarán.

Supongamos en el ejemplo 6.1 que la demanda de las unidades de disco fue más alta de lo pensado. Supongamos que los requerimientos respectivos en las cuatro bodegas son ahora: Amarillo, 90; Teaneck, 78; Chicago, 55, y Sioux Falls, 55. Esto significa que la demanda total es 278 y el abastecimiento total es 260. Para convertir esto en un problema balanceado, agregamos una fábrica ficticia para que represente el déficit de 18 unidades. Esto puede etiquetarse como una fila ficticia en la tabla de transporte, asignando a todas las entradas para esa fila un costo unitario arbitrariamente más grande. La heurística avara puede aplicarse entonces al nuevo cuadro, que tendría cuatro filas y cuatro columnas. No presentaremos los detalles, pero el lector debe verificar que la heurística avara aplicada a este problema desbalanceado envíe todo el déficit a la bodega de Sioux Falls. Observe que cuando el abastecimiento excede la demanda y agregamos una columna ficticia, los costos en la columna ficticia *no* tienen que ser números muy grandes, pero sí los mismos. (De hecho, podemos asignar cero a todos los costos en la columna ficticia.)

El método de agregar una fila o una columna ficticias también funciona cuando encontramos soluciones óptimas mediante programación lineal. En el ejemplo, asignamos un costo de 10^6 a cada celda en la columna ficticia. La hoja de cálculo de Excel y la solución de Solver resultantes (mostradas en la fila etiquetada como “Valor”) aparecen en la figura 6-4. La solución óptima requiere asignar el déficit a dos bodegas: 8 unidades a Chicago y 10 unidades a Sioux Falls.

Los problemas de transporte desbalanceados también pueden formularse como programas lineales, usando restricciones de inequidad. En el ejemplo previo, donde hay exceso de demanda, puede usarse la igualdad en las primeras tres restricciones para estar seguros de que se ha embarcado todo el abastecimiento, y menor o igual que en las últimas cuatro, donde la inactividad justifica el déficit. Al hacerlo, el lector puede revisar que se obtiene la misma solución alcanzada al agregar una fila ficticia. Este método tiene la ventaja de proporcionar un valor exacto para la función objetivo. (En el caso donde el abastecimiento supera la demanda, el principio es el mismo, pero los detalles difieren. Las primeras tres restricciones de abastecimiento se convierten a la forma *mayores o iguales que*, mientras que las cuatro últimas restricciones de demanda siguen siendo restricciones de igualdad. La inactividad en las primeras tres restricciones corresponde al exceso de abastecimiento.)

FIGURA 6-4
Solución del ejemplo 6.1 con demanda en exceso y fila ficticia.

				Solución del ejemplo 6.1 con demanda en exceso y fila ficticia														
Variables	x11	x12	x13	x14	x21	x22	x23	x24	x31	x32	x33	x34	x41	x42	x43	x44	Oper	Valor	RHS
Valor	0	0	0	45	42	78	0	0	48	0	47	0	0	0	8	10			
Obj Func	250	420	380	280	1280	990	1440	1520	1550	1420	1660	1730	1.E+06	1.E+06	1.E+06	1.E+06	Min	2E+07	
st																			
Restricción 1	1	1	1	1													=	45	60
Restricción 2					1	1	1	1									=	120	130
Restricción 3									1	1	1	1					=	95	95
Restricción 4													1	1	1	1	=	18	18
Restricción 5	1				1				1				1				=	90	90
Restricción 6		1				1				1				1			=	78	78
Restricción 7			1				1				1				1		=	55	55
Restricción 8				1				1				1				1	=	55	55

6.4 FORMULACIONES DE REDES MÁS GENERALES

El problema de transporte es un tipo especial de red donde todos los nodos son de abastecimiento (también llamados fuentes) o de demanda (también llamados depósitos). La programación lineal también puede usarse para resolver problemas de distribución de red más complejos. Un ejemplo es el *problema de trasbordo*. En este caso, uno o más de los nodos en la red son puntos de trasbordo más que puntos de demanda o abastecimiento. Observe que el nodo de trasbordo también puede ser un nodo de demanda o de abastecimiento (pero ningún nodo puede ser de abastecimiento y demanda al mismo tiempo).

Para los problemas generales de flujo de red, usamos el siguiente balance de reglas de flujo:

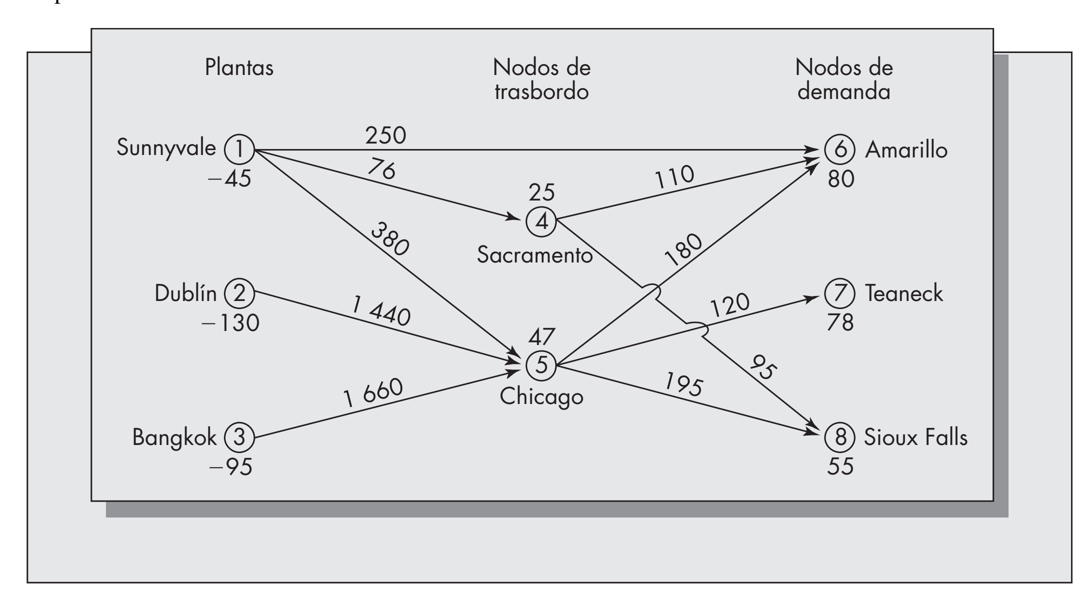
Si	**Aplique la siguiente regla en cada nodo:**
1. Abastecimiento total > demanda total	Flujo de entrada − flujo de salida ≥ abastecimiento o demanda
2. Abastecimiento total < demanda total	Flujo de entrada − flujo de salida ≤ abastecimiento o demanda
3. Abastecimiento total = demanda total	Flujo de entrada − flujo de salida = abastecimiento o demanda

Las variables de decisión se definen de la misma manera que para el problema de la transportación simple. Esto es, x_{ij} representa el flujo total desde el nodo i al nodo j. En el caso de los problemas de red de flujo, representamos el abastecimiento como un número negativo anexo al nodo y la demanda como un número positivo anexo a ese nodo. Esta forma de hacer las cosas junto con las reglas de flujo darán como resultado las ecuaciones correctas de balance de flujo.

Ejemplo *6.2*

Consideremos el ejemplo de Pear Disk Drives. La compañía ha decidido poner una bodega en Sacramento que se usará como un nodo de trasbordo, y ha expandido las instalaciones de Chicago para que también puedan realizarse trasbordos. Supongamos que además de ser nodos de trasbordo, Chicago y Sacramento son también nodos de demanda. La nueva red se muestra en la figura 6-5. Observe que se han eliminado varias de las antiguas rutas en la nueva configuración.

FIGURA 6-5
El problema de Pear Disk Drive con nodos de trasbordo.



Definimos una variable de decisión por cada arco en la red. En este caso, existe un total de 10 variables de decisión. La función objetivo es

$$\text{Minimizar} \quad 250x_{16} + 76x_{14} + 380x_{15} + 1440x_{25} + 1660x_{35} + 110x_{46} + 95x_{48} + 180x_{56} + 120x_{57} + 195x_{58}$$

El abastecimiento total disponible sigue siendo de 260 unidades, pero la demanda es de 285 (debido a 25 adicionales que se solicitaron en Sacramento). Por tanto, esto corresponde al caso 2 de las reglas de flujo en que la demanda total excede el abastecimiento total. Aplicando la regla 2 a cada nodo se obtienen las siguientes ocho restricciones para este problema:

Nodo 1: $-x_{14} - x_{15} - x_{16} \leq -45$

Nodo 2: $-x_{25} \leq -120$

Nodo 3: $-x_{35} \leq -95$

Nodo 4: $x_{14} - x_{46} - x_{48} \leq 25$

Nodo 5: $x_{16} + x_{46} + x_{56} - x_{56} - x_{57} - x_{58} \leq 47$

Nodo 6: $x_{16} + x_{46} + x_{56} \leq 80$

Nodo 7: $x_{57} \leq 78$

Nodo 8: $x_{48} + x_{58} \leq 55.$

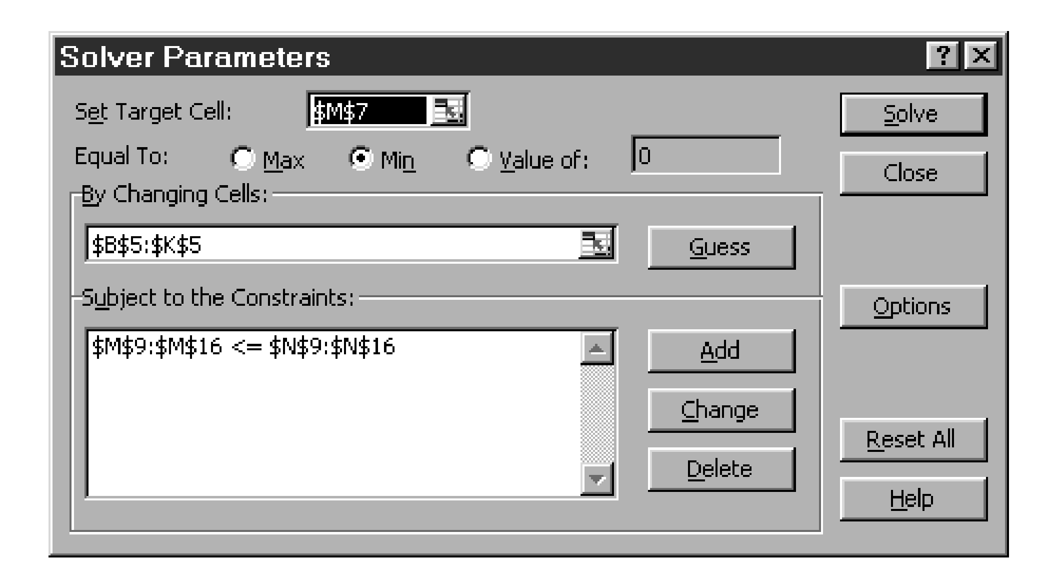
Para algunos códigos de programación lineal, todos los lados derechos deben ser no negativos. En esos casos, deben multiplicarse las primeras tres restricciones por −1. Sin embargo, esto no se requiere en Excel, así que podemos introducir las restricciones tal como aparecen.

La hoja de cálculo de Excel y la solución para el problema de trasbordo de Pear aparece en la figura 6-6. Observe que la solución exige el embarque de todas las unidades desde las fuentes. Como el problema tiene más abastecimiento que demanda, es interesante ver dónde se presentan los déficit. La cantidad enviada a Sacramento es de 45 unidades (todas desde Sunnyvale), y el total enviado desde Sacramento es 20 unidades. Por tanto, en Sacramento se satisface toda la demanda (25 unidades). Para el otro punto de trasbordo, Chicago, los embarques son 120 unidades desde Dublín y 95 unidades desde Bangkok, y los embarques que salen de Chicago son 48 unidades a Amarillo y 120 unidades a Teaneck.

FIGURA 6-6

Hoja de cálculo de Excel para el problema de trasbordo de Pear en el ejemplo 6.2.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		**Hoja de cálculo de Excel para el problema de trasbordo de Pear en el ejemplo 6.2**												
2														
3	Variables	×14	×15	×16	×25	×35	×46	×48	×56	×57	×58	Operador	Valor	RHS
4														
5	Valores	45	0	0	120	95	0	20	80	78	10			
6														
7	Fun obj	76	380	250	1440	1660	110	95	180	120	195	Min	361530	
8	st													
9	Nodo 1	−1	−1	−1								<=	−45	−45
10	Nodo 2				−1							<=	−120	−120
11	Nodo 3					−1						<=	−95	−95
12	Nodo 4	1					−1	−1				<=	25	25
13	Nodo 5		1		1	1			−1	−1	−1	<=	47	47
14	Nodo 6			1			1		1			<=	80	80
15	Nodo 7									1		<=	78	78
16	Nodo 8							1			1	<=	30	55



La diferencia es 120 + 95 − (48 + 120) = 47 unidades. Por tanto, no hay desabasto en Chicago. Los embarques totales en los nodos de demanda en Amarillo, Teaneck y Sioux Falls son respectivamente 80, 78 y 30 unidades. Por tanto, todo el déficit (25 unidades) es absorbido por las instalaciones de Sioux Falls en la solución óptima.

Las redes reales pueden ser extremadamente complejas en virtud de su magnitud total. Como Simchi-Levi *et al.* (1999) dicen, un sistema de distribución normal de refrescos puede involucrar cualquier cosa entre las 10 000 y las 120 000 cuentas. Algunos detallistas tienen miles de tiendas y cientos de miles de productos. Por ello, puede requerirse de una agregación de datos eficiente para resolver problemas de esta magnitud. Por lo general, la agregación del cliente se

logra al combinar cuentas cercanas. Combinar clientes con códigos postales iguales o similares es una forma muy común de agregación geográfica. La agregación de producto se comenta con detalle en el capítulo 3 en el contexto de la fabricación. Las reglas de agregación de producto para una cadena de abastecimiento parecen ser diferentes de las comentadas en el capítulo 3. Por ejemplo, podemos agregar productos conforme al lugar donde van a recogerse o donde se entregan. Recomendamos al lector interesado consultar a Simchi-Levi *et al.* (1999) para un análisis más amplio de las cuestiones prácticas que rodean la implementación de las redes de cadena de abastecimiento.

El modelado matemático se ha utilizado con gran éxito en muchas aplicaciones de cadena de abastecimiento. Los modelos sofisticados son el centro de varios productos de software comercial, como los de i2 Technologies de Texas. La Aplicación instantánea de esta sección trata sobre una exitosa aplicación de modelos avanzados de control de inventario para la cadena de abastecimiento de refacciones de IBM.

Problemas para las secciones 6.1 a 6.4

1. Consideremos el ejemplo 6.1 de Pear Company suponiendo los siguientes abastecimientos y demandas:

Planta	Producción	Bodega	Requerimiento
Sunnyvale	60	Amarillo	100
Dublín	145	Teaneck	84
Bangkok	125	Chicago	77
		Sioux Falls	69

   a) ¿Cuál es la solución obtenida usando la heurística avara?

   b) Utilice Solver de Excel (u otros códigos de programación lineal) para determinar la solución óptima.

   c) ¿Cuál es el error porcentual de la solución que obtuvo en la parte *a*)?

2. Resuelva el problema de transporte de Pear suponiendo que se eliminan las siguientes rutas: de Sunnyvale a Teaneck, de Dublín a Chicago, y de Bangkok a Amarillo. ¿Cuál es el incremento porcentual en los costos de embarque total en la solución óptima por la eliminación de estas rutas?

3. Resuelva el problema de trasbordo de Pear suponiendo que un punto de trasbordo se localiza en la ciudad de Oklahoma. Suponga que los costos unitarios de realizar el embarque desde las tres plantas a Oklahoma son Sunnyvale, \$170; Dublín, \$1 200, y Bangkok, \$1 600, y los costos respectivos de embarcar desde Oklahoma a los nodos de demanda son Amarillo, \$35, Teaneck, \$245, y Sioux Falls, \$145. Suponga que Oklahoma es sólo punto de trasbordo y no tiene demanda para sí misma. Encuentre el nuevo patrón de embarque con la adición del nuevo punto de trasbordo y los ahorros, si los hay, causados por la introducción de este nodo adicional.

4. Major Motors produce su modelo Trans National en tres plantas localizadas en Flint, Michigan; Fresno, California, y Monterrey, México. Los distribuidores reciben coches desde los centros de distribución regional localizados en Phoenix, Arizona; Davenport, Iowa, y Columbia, Carolina del Sur. La producción anticipada en las plantas durante el próximo mes (en cientos de autos) es 43 en Flint, 26 en Fresno y 31 en Monterrey. Basado en las órdenes de la compañía y otras solicitudes de los distribuidores, Major Motors ha decidido que necesita tener las siguientes cantidades en los centros de distribución regional al final del mes:

Aplicación instantánea

IBM HACE MÁS EFICIENTE SU CADENA DE ABASTECIMIENTO PARA REFACCIONES UTILIZANDO MODELOS MATEMÁTICOS SOFISTICADOS

La compañía IBM debe manejar una red muy compleja de bodegas y centros de servicio para proporcionar soporte de refacciones a su base de clientes. El éxito de IBM se ha debido, en parte, a su capacidad para proporcionar un servicio rápido y eficiente a sus clientes. Conforme crecen la base instalada y la variedad de productos, la administración de esta distribución e inventario de sus refacciones se vuelve más compleja. A partir de 1990 IBM tuvo aproximadamente mil diferentes productos en servicio y más de diez millones de éstos en las ubicaciones de usuarios a escala mundial. Se requerían aproximadamente 200 000 partes distintas para dar soporte a estos productos. Debían enviarse desde el fabricante o proveedor a grandes bodegas regionales o nacionales. Desde ahí deben enviarse a bodegas más pequeñas, y después a estaciones locales y finalmente a máquinas instaladas. Un sistema donde las partes se almacenan en diferentes clases de ubicaciones se conoce como sistema de inventarios multinivel. Los escalones se refieren a la jerarquía de los lugares de almacenamiento.

A mediados de los ochenta, la División de Servicio Nacional de IBM (NSD) contactó a un equipo de investigadores de la Escuela Wharton para considerar formas de mejorar el desempeño del sistema de entrega de refacciones mientras se reducen los costos de operación. El equipo creó un modelo matemático sofisticado y, usando una heurística efectiva, fue capaz de proporcionar mejoras significativas en el desempeño del sistema. El suyo desarrolló recomendaciones para establecer las políticas de almacenamiento de inventario para todas las partes en todos los sitios de almacenamiento. Se integró en un sistema más completo que consistía de módulos para contabilidad, distribución central, sistema de tiempo real, planeación y control, y sistemas de campo. Las matemáticas subyacentes usaron conceptos de servicio, costos, niveles de orden, tiempos de inactividad, e incertidumbre de la demanda, comentados en el capítulo 5 de este libro.

¿Cuáles fueron las mejoras observadas después de la implementación del enfoque del equipo? Primero, el optimizador recomendó niveles de inventario que fueron de 20 a 25 por ciento por debajo de los niveles actuales en ese entonces. En segundo lugar, la reducción de inventarios estuvo acompañada de una mejora demostrable en el servicio al cliente y un incremento en la sensibilidad del sistema de control. Entre las medidas concretas tomadas para la mejora podemos mencionar el 10 por ciento de mejora en la disponibilidad de partes en los niveles más bajos en el sistema y ahorros de aproximadamente $20 millones al año.

¿Cuál es la lección que aprendimos? Conforme los sistemas de entrega de productos se hacen cada vez más grandes y complejos, se requieren métodos cada vez más sofisticados para manejar esos sistemas. Los productos de software genérico pueden no ser capaces de proporcionar poder y personalización suficientes para resultar efectivos en dichos ambientes. La experiencia de IBM es sólo un ejemplo de la forma en que la administración de estructuras complejas de cadena de abastecimiento puede mejorar con la ayuda de los métodos de modelado comentados en el presente texto.

Fuente: Cohen *et al*., 1990.

Phoenix, 26; Davenport, 28, y Columbia, 30. Suponga que el costo de embarcar 100 autos desde cada planta hacia cada centro de distribución se da en la siguiente matriz (en miles de dólares):

		A		
		Phoenix	**Davenport**	**Columbia**
DE	**Flint**	12	8	17
	Fresno	7	14	21
	Monterrey	18	22	31

a) Convierta el problema en un problema balanceado agregando una fila o columna apropiada y encuentre la solución sugerida por la heurística avara.

b) Encuentre la solución óptima utilizando Solver y restricciones de inequidad.

c) ¿Cuál es el porcentaje de error de la heurística avara para este problema?

d) Supongamos que la ruta entre Monterrey y Columbia ya no está disponible debido a un desprendimiento de tierra en una de las carreteras principales en México. ¿Cuál solución obtendría ahora?

5. Considere la situación de Major Motors descrita en el problema 4. A fin de poder hacer frente con mayor efectividad a los efectos imprevisibles (como el cierre de carreteras), Major Motors ha establecido dos puntos de trasbordo entre las fábricas y los centros de distribución regional en Santa Fe, Nuevo México y la ciudad de Jefferson en Missouri. El costo de embarcar 100 autos a los puntos de trasbordo sería (en miles de dólares):

		A	
		Santa Fe	**Jefferson City**
DE	**Flint**	8	6
	Fresno	6	9
	Monterrey	9	14

mientras que el costo de realizar un embarque desde los puntos de trasbordo a los centros de distribución es

		A		
		Phoenix	**Davenport**	**Columbia**
DE	**Santa Fe**	3	8	10
	Jefferson City	5	5	9

Suponiendo que ninguna de las rutas directas entre las fábricas y los centros de distribución está disponible, encuentre el flujo óptimo de automóviles a través de los puntos de trasbordo que minimice los costos totales de embarque.

6. Toyco produce una línea de muñecas Bonnie y accesorios en sus plantas de Nueva York y Baltimore, que deben enviarse a los centros de distribución en Chicago y Los Ángeles. La compañía contrató a Air Freight, Inc., para hacer sus embarques. Suponga que puede hacerlo directamente o a través de Pittsburgh y Denver. Las tasas diarias de producción en las plantas son respectivamente 5 000 y 7 000 unidades diariamente y las demandas en los centros de distribución son respectivamente 3500 y 8 000 a diario. Los costos de embarque de mil unidades se dan en la siguiente tabla. Encuentre las rutas óptimas de embarque y el costo asociado.

		A			
		Pittsburgh	**Denver**	**Chicago**	**Los Ángeles**
DE	**Nueva York**	$182	$375	$285	$460
	Baltimore	77	290	245	575
	Pittsburgh	—	275	125	380
	Denver	—	—	90	110

7. Reconsidere el problema 6 si hay una caída en la demanda de muñecas a 3 000 en Chicago y 7 000 en Los Ángeles. Encuentre el patrón óptimo de embarque en este caso. ¿Qué tanto de la disminución total de 2 000 unidades es absorbida en cada fábrica en la solución óptima?
8. Reconsidere el problema 6 suponiendo que la cantidad máxima que puede embarcarse bien sea desde Nueva York o Baltimore a través de Pittsburgh es de 2 000 unidades debido al tamaño del plano disponible para esta ruta.

6.5 PLANEACIÓN DE RECURSOS DE DISTRIBUCIÓN

La planeación de recursos de distribución (DRP, *distribution resource planning*) es un término acuñado por Martin (1990). En corto, se utiliza la lógica MRP desarrollada para los sistemas de fabricación en el manejo de los sistemas de distribución. La principal ventaja de DRP sobre el ROP tradicional (sistemas de punto de reorden, comentados en el capítulo 5) es que pueden enfrentar patrones cambiantes de demanda. Sin embargo, mientras el ROP incorpora claramente la incertidumbre en la demanda (es decir, el error de pronóstico), la DRP no lo hace. La DRP es muy parecida a MRP en cuanto a su concepto y ejecución. Debido a que MRP se comentará en detalle en el capítulo 7, proporcionaremos sólo un breve panorama general del concepto.

Ejemplo ***6.3***

La librería de la universidad ha decidido implementar un enfoque DRP para los nuevos pedidos de guías, ya que se ha observado amplia variación en la demanda. Las ventas tienden a elevarse justo antes de los periodos de exámenes y son muy bajas en las vacaciones de invierno y primavera. Las guías se embarcan en paquetes de 20. Supongamos que las ventas pronosticadas durante las siguientes ocho semanas (medidas en paquetes) son 22, 35, 60, 12, 0, 19, 85, 33. El proveedor ha embarcado 40 paquetes, que están en tránsito y esperan ser entregados. Más aún, actualmente hay 26 paquetes pedidos y programados para llegar en dos semanas y 26 paquetes en inventario. Actualmente, el perfil de pedido DRP de guías es

Perfil de pedidos basados en DRP para guías

Existencias	26
Demora de la orden	2 semanas
Inventario de seguridad	0

Semanas:	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Pronósticos de ventas		22	35	60	12	0	19	85	33
En tránsito		40							
Ordenados			26						
Balance proyectado	26	44	35	−25					

El perfil DRP muestra que, sin ninguna otra acción, la librería se desprenderá de 25 paquetes en la semana 3. Para evitar este desabasto, es necesario hacer un pedido antes de la semana 3 (a saber, en la semana 1).

Al representar el perfil de inventario de esta manera, la magnitud y distribución del tiempo de desabastos quedarán claras antes de ocurrir. Cuando se utiliza el DRP no se pretende hacer pedidos con demasiadas semanas a futuro, ya que los pronósticos pueden cambiar conforme llega nueva información. Más aún, esperamos también que las ventas reales sean diferentes de las proyectadas. La representación clara del historial del pedido, llegadas y desabastos anticipados tiene ventajas sobre la presentación de datos tradicional del ROP. Según Martin (1990):

> Obviamente, el modelo de inventario DRP es muy superior. … Los datos cobraron vida, cuentan una historia y demandan acción. Proyectan y anticipan desabastos en el futuro y recomiendan acción

sugiriendo cantidades de pedido por fases de tiempo. El resultado es que la actividad de compra no es ni prematura ni retrasada, sino justo a tiempo y en las cantidades adecuadas. Conforme varían las condiciones (y lo harán), este enfoque de modelado de inventario se actualiza a sí mismo de forma muy dinámica, identifica los desabastos potenciales futuros y envía adecuados mensajes de acción a sus usuarios.

Ejemplo *6.3* *(continuación)*

Continuemos con el ejemplo de la librería. En este punto, queda claro que un pedido necesita hacerse en la semana 1 para evitar el desabasto en la semana 3. Esto se muestra en la siguiente tabla.

Perfil de pedidos basados en DRP para guías

Existencia	26
Tiempo de demora	2 semanas
Inventario de seguridad	0

Semanas:	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Pronósticos de ventas		22	35	60	12	0	19	85	33
En tránsito		40							
En pedido			26						
Balance proyectado	26	44	35	12	0	0	−19		
Embarques planeados fecha de pedido		37							
Embarques planeados fecha de recepción				37					

La tabla muestra un embarque de 37 paquetes programados para llegar en la semana 3. Esto evita el desabasto anticipado en esta semana y también la necesidad de que la librería deba hacer otro pedido de nuevo hasta la semana 4 (para impedir el desabasto anticipado en la semana 6).

Aunque Martin no comenta al respecto, podemos preguntarnos qué tan grandes deben ser los embarques planeados. Si vemos los pronósticos como una corriente determinista de las demandas, entonces el problema al determinar los tamaños óptimos de pedido es justo el mismo que el tratado en el capítulo 7, y consiste en determinar los tamaños de lote de producción óptimos en MRP. En el capítulo 7 comentamos varios métodos para determinar el tamaño de lote, que también se aplican en el contexto DRP: el método Wagner-Whitin, balance de periodo parcial, y la heurística Silver-Meal. Para aplicar cualquiera de estos métodos, es necesario conocer los costos de mantener el inventario y los de realizar el pedido.

Aunque no aparece en este ejemplo, el enfoque DRP puede incorporar positivas existencias de seguridad. Los pedidos deben planearse de manera que el inventario existente anticipado jamás caiga por debajo del abastecimiento de seguridad. Por otra parte, podemos reemplazar la demanda pronosticada con la demanda pronosticada más el inventario de seguridad (es decir, "demanda efectiva") y determinar los tamaños de pedido para cumplir con la demanda efectiva. Sin embargo, la DRP no recomienda niveles de existencias de seguridad. El cálculo apropiado de niveles de existencias de seguridad requiere de una evaluación de la incertidumbre de demanda y un cálculo de costos de desabasto o especificación de los niveles de servicio. Este problema se comentó en detalle en el capítulo 5.

Aunque sencilla, la DRP puede ser una herramienta efectiva, en especial para los sistemas complejos de distribución multinivel. Dichos sistemas normalmente incluyen una o más capas de centros de distribución intermedios. Observe que existe una gran cantidad de literatura teórica en modelos matemáticos para manejar sistemas de inventario multinivel (conocidos como sistemas multiescalón). Sin embargo, la mayoría de estos modelos son matemáticamente muy complejos y raras veces resultan adecuados para la implementación en sistemas de gran escala.

Problemas para la sección 6.5

9. Considere el ejemplo 6.3 de esta sección. Con base en las ventas pronosticadas para las siguientes ocho semanas dadas, determine las ventas promedio en este periodo. Suponga que el costo fijo de hacer un pedido es $40 y el costo de mantener el inventario por paquete es $0.25. Utilice los números para determinar un valor para CEP (véase el capítulo 4). Suponiendo que el CEP se ordena de forma tal que los inventarios jamás caigan por debajo de cero, determine el pedido DRP de guías durante las siguientes ocho semanas. (Considere las cifras en tránsito y en existencia dadas en el problema.)
10. Utilizando los datos del problema 9, determine el tamaño de lote con base en la heurística Silver-Meal (comentada en el capítulo 7) y muestre el perfil de pedido DRP en este caso.
11. Resuelva de nuevo el problema 10 suponiendo que desea mantener existencias de seguridad de 30 unidades, y muestre el perfil DRP en este caso.
12. Comente las ventajas y desventajas de la DRP en comparación con el ROP (sistemas de punto de reorden), comentado en detalle en el capítulo 5.

6.6 DETERMINACIÓN DE LAS RUTAS DE ENTREGA EN LAS CADENAS DE ABASTECIMIENTO

Un importante aspecto de la logística de la cadena de abastecimiento es mover con eficiencia el producto de un lugar a otro. Los problemas de transporte y trasbordo que se analizaron antes en este capítulo abordan este problema en un nivel macro o de toda la compañía. En un nivel micro, las entregas a los clientes también deben planearse con mucha eficiencia. Debido a su escala, los problemas de una entrega eficiente tienen un impacto significativo en la base. Como resultado, se convierten en una parte importante del diseño de toda la cadena de abastecimiento.

En general la determinación de programas óptimos de entrega resulta un problema muy difícil, tanto como los problemas de programación de taller analizados en detalle en el capítulo 8. La programación de vehículos está muy relacionada con un problema clásico de investigación de operaciones conocido como el problema del agente viajero. Se describe de la siguiente forma: un vendedor empieza en su lugar de residencia, etiquetado ciudad 1. Debe entonces hacer paradas en $n - 1$ otras ciudades, visitando cada ciudad una sola vez. El problema es determinar la secuencia óptima en la cual hay que visitar las ciudades para minimizar la distancia total recorrida. Aunque este problema es fácil de establecer, resolverlo es muy difícil. Si el número de ciudades es pequeño, se pueden enumerar todos los viajes posibles. Existen $n!$ pedidos de objetos.[1] Para valores pequeños de n, pueden enumerarse todos los viajes y calcular directamente las distancias. Por ejemplo, para $n = 5$ hay 120 secuencias. Sin embargo, este número crece muy rápido. Para $n = 10$, el número de secuencias crece a más de 3 millones, y para $n = 25$ crece más de 1.55×10^{25}. Para tener cierta idea de qué tan grande es este número, supongamos que podemos evaluar un billón de secuencias por segundo en una supercomputadora. Entonces, para un problema de 25 ciudades, se necesitarían casi 500 000 años para evaluar todas las secuencias.

La enumeración total es imposible para resolver otra cosa que no sean los problemas más pequeños del agente viajero. Los problemas como éstos se conocen en matemáticas como *NP duro*. Con NP se hace referencia a no polinomial, es decir, que el tiempo requerido para resolver dicho problema es una función exponencial del número de ciudades más que una función polinomial. No podemos explayarnos en el problema del agente viajero, pero observemos que los métodos de solución propuestos son grandes mejoras sobre la enumeración total. No obstante, sigue siendo difícil encontrar soluciones óptimas incluso para problemas de tamaño mediano.

---

[1] $n!$ es igual a n veces $(n - 1)$ veces $(n - 2)$ . . . veces 1.

La determinación de rutas óptimas en la programación de vehículos es un problema similar, pero más complejo. Supongamos que existe un depósito central con uno o más vehículos de entrega y n ubicaciones de clientes, cada una de ellas con un requerimiento conocido. La pregunta es cómo asignar los vehículos a las instalaciones de los clientes para cubrir sus demandas y satisfacer cualesquiera restricciones que pudieran tener a un costo mínimo. Los problemas de programación de vehículos más reales son demasiado grandes y complejos para tener una solución óptima.

Debido a que la optimización puede ser imposible de alcanzar, es importante contar con métodos para determinar "buenas" soluciones. Comentaremos una sencilla técnica para encontrar buenas rutas, conocida como el método de ahorros, desarrollada por Clarke y Wright (1964).

Supongamos que existe un solo depósito desde el cual entran y salen todos los vehículos. Se conocen las ubicaciones y necesidades de los clientes. Identifiquemos el depósito como ubicación 0 y a los clientes como ubicaciones 1, 2, . . . , n. Suponemos que existen costos conocidos por viajes del depósito a las instalaciones de cada cliente, dados por

$$c_{0j} = \text{Costo de realizar un viaje de un depósito al cliente } j.$$

Para implementar el método, necesitamos conocer también los costos de viajes entre clientes. Significa que supondremos que se conocen también las siguientes constantes:

$$c_{ij} = \text{Costo de realizar un viaje desde la ubicación de un cliente } i \text{ a la ubicación de cliente } j.$$

Para nuestros fines consideramos sólo el caso en que $c_{ij} = c_{ji}$ para todos $1 \leq i, j \leq n$. Esto no se mantiene necesariamente en todas las situaciones. Por ejemplo, si hay calles de un solo sentido, la distancia de i a j puede ser diferente a la distancia de j a i. El método procede como sigue: supongamos que, de entrada, existe un vehículo separado asignado a cada ubicación del cliente. La solución inicial consiste en n rutas separadas desde el depósito hasta las instalaciones del cliente y de regreso. Se deduce que el costo total de realizar todos los viajes redondos para la solución inicial es:

$$2 \sum_{j=1}^{n} c_{0j}.$$

Ahora, supongamos que relacionamos los clientes i y j. Esto es, vamos del depósito a i a j y de vuelta al depósito. Al hacer esto, podemos ahorrarnos un viaje entre el depósito y la ubicación i y un viaje entre el depósito y la ubicación j. Sin embargo, habrá un costo añadido de c_{ij} para el viaje de i a j (o viceversa). Por tanto, los ahorros realizados al unir i y j son

$$s_{ij} = c_{0i} + c_{0j} - c_{ij}.$$

El método consiste en calcular s_{ij} para todos los posibles pares de ubicaciones de clientes y después clasificar el s_{ij} en orden decreciente. Podemos considerar entonces cada uno de los enlaces en orden descendente de ahorros e incluir el enlace (i, j) en una ruta siempre y cuando no viole las restricciones de factibilidad. Si al incluir el enlace actual se viola la factibilidad, pasamos al siguiente enlace en la lista y consideramos incluirlo en una sola ruta. Continuamos de esta manera hasta que se agote la lista. Cuando el enlace (i, j) se incluyó en una ruta, el ahorro de costos es s_{ij}.

El número total de cálculos de s_{ij} que se requiere es

$$\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2}.$$

(Cuando c_{ij} y c_{ji} no son iguales, es necesario calcular dos veces los términos de ahorro.)

El método de ahorros es factible para resolverse a mano sólo en caso de valores pequeños de n. Por ejemplo, para $n = 10$ existen 45 términos, y para $n = 100$ existen casi 5 000 términos. Sin embargo, siempre que las restricciones no sean demasiado complejas, el método puede implementarse fácilmente en una computadora.

Ilustramos el método con el siguiente ejemplo.

Ejemplo *6.4*

Whole Grains es una pequeña panadería que vende cada mañana a cinco clientes principales. Si la ubicamos en el punto de origen de una cuadrícula [es decir, en el punto (0, 0)], entonces la ubicación de los cinco clientes principales y sus requerimientos diarios son

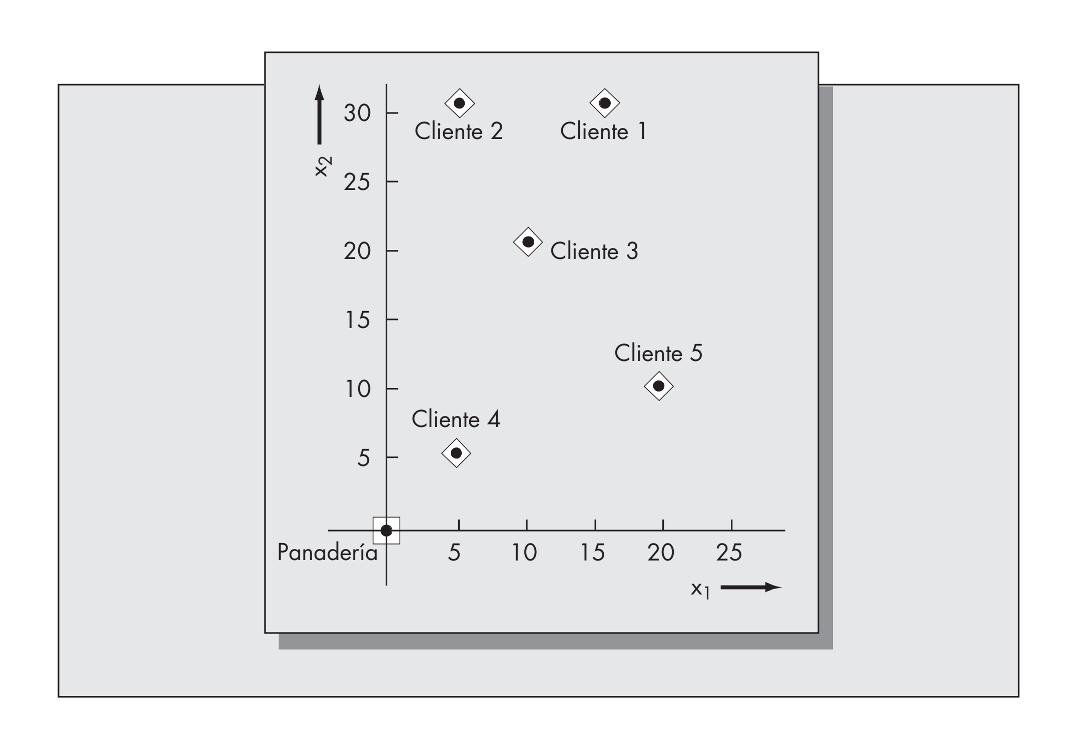
Cliente	**Ubicación**	**Requerimientos diarios (hogazas)**
1	(15, 30)	85
2	(5, 30)	162
3	(10, 20)	26
4	(5, 5)	140
5	(20, 10)	110

Las ubicaciones relativas de Whole Grain y sus cinco clientes se muestran en la figura 6-7. La panadería tiene varios camiones de entrega, cada uno de los cuales tiene una capacidad de 300 hogazas. Debemos asumir que el costo de viaje entre dos ubicaciones cualesquiera es simplemente la línea recta o distancia euclidiana entre los puntos. La fórmula para la distancia de línea recta que separa los puntos (x_1, y_1) y (x_2, y_2) es

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

La meta es encontrar un patrón de entrega que cumpla con la demanda del cliente y al mismo tiempo minimice los costos de entrega, sin exceder la capacidad según el tamaño de los camiones de entrega.

FIGURA 6-7
Ubicaciones de los clientes en el ejemplo 6.4.



Solución

El primer paso es calcular el costo de cada par (i, j) donde i y j pueden variar de 0 a 5. Estamos suponiendo que este costo es la distancia de línea recta entre los puntos que representan las ubicaciones del cliente. Las distancias en línea recta se dan en la matriz siguiente.

Matriz de costos (c_{ij})

		A					
		0	**1**	**2**	**3**	**4**	**5**
DE	**0**		33.5	30.4	22.4	7.1	22.4
	1			10.0	11.2	26.9	20.6
	2				11.2	25.0	25.0
	3					15.8	14.1
	4						15.8

Después, calculamos los ahorros de todos los pares (i, j), $1 \leq i < j \leq 5$. Existe un total de 10 términos de ahorro a calcular para este ejemplo:

$$s_{12} = c_{01} + c_{02} - c_{12} = 33.5 + 30.4 - 10 = 53.9,$$

$$s_{13} = c_{01} + c_{03} - c_{13} = 33.5 + 22.4 - 11.2 = 44.7.$$

Los términos restantes se calculan de la misma manera, con los resultados

$$s_{14} = 13.7, \quad s_{25} = 27.8,$$

$$s_{15} = 35.3, \quad s_{34} = 13.7,$$

$$s_{23} = 41.6, \quad s_{35} = 30.7,$$

$$s_{24} = 12.5, \quad s_{45} = 13.7.$$

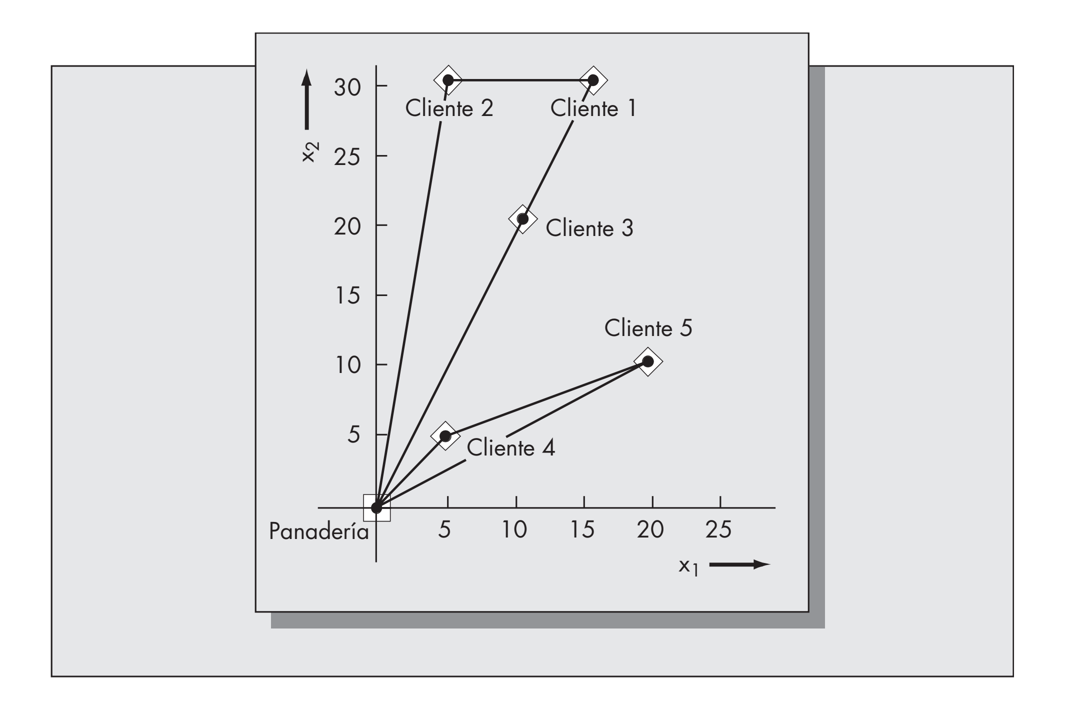
El siguiente paso es clasificar los pares de clientes en orden decreciente de sus valores de ahorro. Esto da como resultado la clasificación

(1, 2), (1, 3), (2, 3), (1, 5), (3, 5), (2, 5), (1, 4), (3, 4), (4, 5) y (2, 4).

Observe que (1, 4), (3, 4) y (4, 5) tienen los mismos ahorros. Los lazos se rompen arbitrariamente, así que estos pares pueden clasificarse de forma diferente. Ahora comenzamos a combinar clientes y a crear rutas de vehículos al considerar los pares en orden de clasificación, revisando cada vez no violar las restricciones del problema. Como (1, 2) es el primero en la lista, de entrada intentamos relacionar a los clientes 1 y 2 en la misma ruta. Hacer esto da como resultado una carga de 85 + 162 = 247 hogazas. Después, consideramos combinar 1 y 3, lo que significa incluir 3 en la misma ruta. Esto da como resultado una carga de 247 + 26 = 273, que continúa siendo factible. Por tanto, hemos construido ahora una ruta consistente de los clientes 1, 2 y 3. El siguiente par en la lista es (2, 3). Sin embargo, 2 y 3 ya están en la misma ruta. El que sigue en la lista es (1, 5). No obstante, no es factible relacionar al cliente 5 con la ruta actual. Debido a que la demanda de la ubicación 5 es de 110 hogazas, añadir la ubicación 5 a la ruta actual excederá la capacidad del camión. El siguiente par factible de la lista es (4, 5) que hacemos en una nueva ruta. La solución recomendada por el método de ahorros consiste en dos rutas, como se muestra en la figura 6-8.

Debemos señalar que el método de ahorros es sólo una heurística. No necesariamente produce una ruta óptima. El problema es que al forzar la opción de un enlace clasificado alto se pueden excluir otros enlaces que podrían tener ahorros ligeramente menores pero quizás resulten mejores opciones en un sentido global, permitiendo que se elijan otros enlaces hacia abajo. Varios autores han sugerido modificaciones del método de ahorros para intentar superar esta dificultad. El lector interesado puede consultar a Eilon *et al*. (1971) para ver un análisis sobre estos métodos. De cualquier manera, los autores señalan que estas modificaciones no siempre resultan una solución más efectiva en cuanto al costo.

FIGURA 6-8
Ruta de vehículos encontrada a partir del método de ahorros para el ejemplo 6.4.



Aspectos prácticos en la programación de vehículos

Podemos clasificar los problemas de programación de vehículos como uno de dos tipos: basados en los arcos y basados en los nodos. Los primeros son aquellos donde la meta es cubrir una cierta colección de arcos en una red. Ejemplos típicos son la remoción de nieve y la recolección de basura. El tipo de problemas de distribución que hemos comentado en esta sección son los basados en los nodos. El objetivo es visitar un conjunto específico de ubicaciones. Los problemas pueden ser una combinación de ambos.

Los problemas reales de programación de vehículos son mucho más complejos que los que se describen en el ejemplo 6.4. Schrage (1981) enumera 10 características que hacen que los problemas reales sean difíciles de resolver. Incluyen las siete características siguientes:

1. *Requerimientos de frecuencia.* Es posible que las visitas a los clientes deban realizarse con cierta frecuencia, la cual puede variar de un cliente a otro. En nuestro ejemplo, las entregas de pan se realizan a diario, por lo que la frecuencia no representa ningún problema. Consideremos, sin embargo, el problema de entregar gas o petróleo para uso residencial. La frecuencia de entrega depende del uso, así que la frecuencia de entrega variará de un cliente a otro.

2. *Ventanas de tiempo.* Esto se refiere al requerimiento de que las visitas a las ubicaciones de los clientes se hagan en tiempos específicos. Los sistemas de entrega a domicilio y las recolecciones y entregas bancarias y postales son algunos ejemplos comunes.

3. *Tiempo de viaje dependiente del tiempo.* Cuando las entregas se realizan en centros urbanos, un factor que hay que tomar en cuenta es la carga de tráfico durante las horas pico. Éste es un ejemplo del caso en el que el tiempo de viaje (y por tanto el costo asociado con un enlace en la red) depende de la hora del día.

4. *Restricciones de capacidad multidimensional.* Existen muchas restricciones en el peso, así como en el volumen. Esto puede resultar un asunto espinoso, en especial cuando los mismos vehículos se utilizan para transportar diversos productos diferentes.

Aplicación instantánea

AIR PRODUCTS LOGRA IMPORTANTES AHORROS CON EL OPTIMIZADOR DE RUTAS Y PROGRAMACIÓN

Air Products Corporation es un productor de gases industriales ubicado en Allentown, Pennsylvania. Sus productos incluyen oxígeno líquido (para producción de acero y saneamiento), y nitrógeno líquido (para congelación rápida de alimentos y operaciones de esmerilado) entre otros. Air Products siempre ha sido líder en el campo y ha marcado nuevos rumbos mediante innovadoras estrategias de distribución, como la construcción de instalaciones cercanas a los clientes grandes para evitar el costoso y peligroso movimiento de los gases. En los setenta, las plantas, camiones y demás equipo utilizado estaba muy automatizado. Esto contrastaba con el sistema manual de programación de entrega de sus productos. Durante este periodo, la gerencia reconoció que a fin de que la compañía se mantuviera en niveles competitivos, debía mejorar su sistema de entregas. La programación de la entrega de gases industriales tiene varias características que hacen el problema más complejo que los tipos de programación de vehículos comentados anteriormente en esta sección. Estas complicaciones añadidas incluyen:

- Las entregas individuales se pueden dividir entre varios camiones, puesto que lo único importante es la cantidad de entrega total.
- Las entregas deben hacerse únicamente dentro de una ventana de tiempo. Los clientes pueden aceptar entregas sólo en ciertos días y en ciertos tiempos.
- Los camiones difieren en capacidades y costos de operación. Es más, las capacidades son obligatorias por la ley estatal y pueden variar de un estado a otro.
- Algunos camiones no pueden atender a algunos clientes. Quizá sean demasiado grandes o se requiera energía externa.

Bell *et al*. (1983) reportan un optimizador de ruta y programación complejo desarrollado para Air Products. Los calendarios se encuentran resolviendo un programa entero-mixto grande en extremo (hasta 800 000 variables y 200 000 restricciones), hasta un punto cercano a la optimalidad, utilizando un método conocido como relajación lagrangiana. Notifican significativos ahorros para Air Products con este sistema. El beneficio más tangible fue la reducción en los gastos de entrega resultantes del incremento en la productividad de los choferes y los vehículos. Se calcula que los ahorros en kilometraje fueron cercanos al 10 por ciento de los costos, lo que se tradujo en millones de dólares para un subsistema de 16 depósitos. Esta aplicación muestra que es posible implementar sofisticadas técnicas de investigación de operaciones en un ambiente real.

5. *Tipos de vehículos.* Las compañías grandes tienen varios tipos de vehículos de donde elegir. Los tipos de vehículos pueden diferir en capacidad, costo de operación y si están restringidos a viajes cerrados (donde deben regresar al depósito después de realizar las entregas). Cuando hay varios tipos de vehículos, el número de alternativas factibles se eleva considerablemente.
6. *Entregas divididas.* Si un cliente tiene un requerimiento particular, puede tener sentido contar con más de un vehículo asignado a ese cliente.
7. *Incertidumbre.* Los algoritmos de ruta suponen invariablemente que toda la información se conoce por adelantado. Sin embargo, en la realidad el tiempo requerido para cruzar ciertas porciones de una red puede ser muy variable, dependiendo de factores como las condiciones de tráfico, el clima y las descomposturas del vehículo.

Problemas para la sección 6.6

13. Resuelva el ejemplo 6.4 suponiendo que la capacidad del vehículo es de sólo 250 hogazas de pan.
14. Resuelva el ejemplo 6.4 suponiendo que la distancia entre las dos ubicaciones es la distancia rectangular en lugar de la distancia euclidiana. (Véase la sección 10.7 para conocer una definición de la distancia rectangular.)

15. Agregue las siguientes ubicaciones y requerimientos de los clientes al ejemplo 6.4 y resuelva:

Cliente	Ubicación	Requerimiento diario
6	(12, 12)	78
7	(23, 3)	126

16. Suponga que se desea programar vehículos de un depósito central a las instalaciones de cinco clientes. El costo de realizar viajes entre cada par de ubicaciones se presenta en la siguiente matriz. (Suponga que el depósito es la ubicación 0.)

Matriz de costos (c_{ij})

		A					
		0	**1**	**2**	**3**	**4**	**5**
	0		20	75	33	10	30
	1			35	5	20	15
D	**2**				18	58	42
E	**3**					40	20
	4						25

Suponga que estos costos corresponden a distancias entre ubicaciones y que cada vehículo está restringido a viajar un máximo de 80 kilómetros en cada ruta. Encuentre la ruta sugerida por el método de ahorros.

17. All Weather-Oil and Gas Company está planeando rutas de entrega a seis clientes de gas natural. Las ubicaciones y requerimientos de gas (en galones) de los clientes se muestran en la siguiente tabla.

Cliente	Ubicación	Requerimientos (galones)
1	(5, 14)	550
2	(10, 25)	400
3	(3, 30)	650
4	(35, 12)	250
5	(10, 7)	300

Suponga que el depósito se localiza en el origen de la cuadrícula y que los camiones de entrega tienen capacidad de 1 200 galones. También, que el costo de viaje entre cualquiera de las dos ubicaciones es la distancia en línea recta (euclidiana) entre ellas. Encuentre el programa de ruta obtenido del método de ahorros.

6.7 DISEÑO DE PRODUCTOS PARA LA EFICIENCIA DE LA CADENA DE ABASTECIMIENTO

El diseño de producto fue tradicionalmente una función completamente divorciada de las cuestiones más mundanas de la administración de operaciones. Los diseñadores se preocupaban más por la estética y la comerciabilidad. En la etapa de diseño poco se atendía a las cuestiones

básicas como la fabricación y la logística. Conforme la calidad y confiabilidad fueron colocándose a la vanguardia, quedó claro que la confiabilidad y el diseño de producto están muy relacionados. El movimiento del diseño para la manufacturabilidad (DFM, *design for manufacturability*) surgió de una necesidad de conocer la razón por la que los productos fallaban y cómo podrían minimizarse dichas fallas. Una vez entendido que la confiabilidad puede evaluarse en la etapa de diseño, se forjó el vínculo entre el diseño y la fabricación. Otro término para el DFM es ingeniería concurrente.

En años recientes las compañías se percataron de que la logística de manejar la cadena de abastecimiento tiene tanto impacto en la base como en el diseño y la fabricación del producto. Más que nunca, vemos diseños innovadores que toman en cuenta las consideraciones de la cadena de abastecimiento. Una de las formas en que se describe este concepto es el diseño de la logística (DEL). Otra es la ingeniería concurrente tridimensional (3-DCE), un término adoptado por Fine (1998). Las tres dimensiones aquí son producto, proceso y cadenas de abastecimiento. El concepto de ingeniería concurrente va un paso adelante. Ingeniería concurrente significa que las cuestiones relacionadas con el producto (funcionalidad, comerciabilidad) y cuestiones relacionadas con el proceso (cómo se genera el producto, calidad y confiabilidad del producto final) son consideraciones conjuntas en la fase de diseño. Ingeniería concurrente tridimensional significa que la logística de la cadena de abastecimiento se considera también en la fase de diseño del producto.

Las dos formas significativas en que las consideraciones de logística ingresan en la fase de diseño de producto son

1. Diseño de producto para el embarque y la transportación eficientes.
2. Aplazamiento de la configuración de producto final (también llamado diferenciación retardada).

Los productos que pueden empacarse, moverse y almacenarse simplifican fácilmente tanto la fabricación como la logística. Los compradores prefieren productos fáciles de almacenar y de mover. Algunos tienden a ser grandes y voluminosos y presentan un desafío especial en este aspecto. Un ejemplo es el mobiliario. La compañía sueca Ikea realizó ciertamente un excelente trabajo en el diseño de productos modulares y fáciles de almacenar. El mobiliario se vende en juegos fáciles de ensamblar, lo que permite a los minoristas de Ikea almacenarlo en la misma bodega de los lugares en que se exhiben. Simchi-Levi *et al.* (1999) comentan otros ejemplos de productos cuyo éxito se basa en parte en su facilidad de embarque y almacenaje.

Aplazamiento en las cadenas de abastecimiento

Uno de los hallazgos clave de la investigación de la cadena de abastecimiento en los años recientes es el valor de aplazar la configuración del producto final tanto como sea posible. La primera aplicación de este principio hasta donde el autor sabe, fue implementada por Benetton (Signorelli y Heskett, 1984). El Grupo Benetton es un fabricante de ropa italiano que para 1982 se había convertido en el líder mundial en el campo de la ropa tejida. Cerca de 60 por ciento de los atuendos vendidos por la firma están hechos de lana. Tradicionalmente la lana se teñía antes de ser tejida. En 1972 Benetton emprendió una estrategia única: teñir los atuendos *después* de haberlos tejido. Podría cuestionarse esta estrategia dado que los costos de producción y mano de obra para atuendos teñidos después de la fabricación son 10 por ciento más altos que para los realizados con hilo previamente teñido.

La ventaja de dar marcha atrás al orden de las operaciones de teñido y tejido es que proporciona tiempo adicional antes de comprometerse a la mezcla final de colores. Este tiempo da a la compañía la posibilidad de obtener datos adicionales en preferencias al consumidor respecto a los colores. La ropa tejida de Benetton incluyó cerca de 500 combinaciones de estilo y color. Las prendas sin teñir se conocen como *existencias grises*. Mantener inventarios de existencias grises en vez de prendas teñidas tiene varias ventajas. En primer lugar, si un color específico se vuelve más popular de lo anticipado, Benetton puede cubrir la demanda de ese color. Segundo, la compañía correrá menos riesgo de tener provisiones sin vender de prendas de colores poco

populares. Estas ventajas compensan con creces los costos extra de teñir las prendas tejidas, en vez de la lana cruda.

¿Cómo puede correlacionarse el aplazamiento con la teoría de administración de inventarios? Según vimos en el capítulo 5, las existencias de seguridad se retienen como protección contra la incertidumbre en la demanda sobre el tiempo inactivo de reabastecimiento. El tiempo inactivo de las prendas de un color específico se reduce posponiendo la operación de teñido. Se deduce que la incertidumbre también baja, logrando así niveles de servicio comparables con menores existencias de seguridad.

El aplazamiento se ha convertido en una estrategia clave en muchas industrias. Otro ejemplo comentado en la literatura es el de Hewlett-Packard (HP) (Lee, Billington y Carter, 1994). HP es uno de los productores líderes de impresoras láser y de inyección de tinta, entre otros productos. Éstas se venden internacionalmente. Mientras que los mecanismos básicos de impresoras vendidas en el extranjero son iguales a las versiones estadounidenses, algunos subensambles como el suministro de electricidad deben personalizarse para los mercados locales. La estrategia original era configurar impresoras para los requerimientos locales (es decir, localizarlos) en la fábrica. Esto es, impresoras con los manuales, enchufes y suministros eléctricos serían producidos en la fábrica, clasificados y enviados al extranjero como productos finales. El resultado fue que HP necesitaba llevar grandes existencias de seguridad de todas las configuraciones de impresoras.

Para reducir inventarios y mejorar el servicio proporcionado por los centros de distribución (CD) a los clientes minoristas, HP adoptó una estrategia similar a la de Benetton. Las impresoras provenientes de la fábrica serían genéricas o existencias grises. La localización se haría en el nivel CD más que en la fábrica. Al igual que para Benetton, esto redujo las existencias de seguridad que antes se necesitaban, al tiempo que mejoró el servicio. Las impresoras se embarcan a centros de distribución en el extranjero por barco, requiriendo un tiempo de tránsito de un mes. Con la personalización local del producto se redujo espectacularmente el tiempo de inactividad de reabastecimiento para las impresoras configuradas localmente. También, la demanda en la fábrica es ahora la totalidad de las demandas en los CD. Como vimos en los capítulos 2 y 5, las demandas totales tienen una variación relativamente menor. Lee, Billington y Carter mostraron que la localización de CD para las impresoras HP Deskjet-Plus conduciría a reducir 18 por ciento en inventarios sin reducción en los niveles de servicio. De acuerdo con Harold E. Edmonson, un vicepresidente de HP (según aparece citado en Lee, Billington y Carter, 1994):

> Los resultados de este análisis de modelo confirmaron la efectividad de la estrategia para localizar las impresoras en centros de distribución remotos. Una estrategia de diseño como ésta tiene significativos beneficios en términos de una mayor flexibilidad para cumplir con las demandas de los clientes, así como ahorros tanto en inventarios como en costos de transportación. Debo agregar que el diseño para el concepto de localización ahora forma parte de nuestra estrategia de distribución y fabricación.

La idea de posponer la personalización de un producto parece ganar aceptación. Por ejemplo, cuando es posible los semiconductores se producen en forma genérica, y se personalizan a través de la programación o vía otros medios después de haberse recibido los pedidos (Barrone, 1996). Las compañías en muchas industrias están tomando conciencia de los beneficios de reducir tiempos inactivos y acumulación de riesgos mediante el aplazamiento y la personalización local de los productos.

Problemas adicionales en el diseño de la cadena de abastecimiento

Aunque los productos pueden diseñarse mejor para una operación eficiente de la cadena de abastecimiento, hay varios aspectos importantes que deben considerarse en el diseño de la cadena de abastecimiento en sí. Fine (1998) se refiere al diseño de la cadena de abastecimiento como la "competencia clave definitiva". Tres aspectos importantes y relevantes son

- La configuración de la base de proveedores.
- Arreglos para subcontratación.
- Canales de distribución.

Configuración de la base de proveedores

El número de proveedores, sus ubicaciones y tamaños son consideraciones importantes para el diseño eficiente de la cadena de abastecimiento. En años recientes, la tendencia ha sido reducir el número de proveedores y desarrollar arreglos a largo plazo con la base de proveedores existente. Un ejemplo es Xerox Corporation. Jacobson y Hillkirk (1986) comentan varias razones para el impresionante cambio de Xerox durante los años ochenta. Una de las estrategias de la compañía era mejorar la cadena de abastecimiento en la parte de las adquisiciones al reducir el número de sus proveedores de 5 000 a sólo 400. Entre ellos se incluyó una mezcla tanto de locales como extranjeros. Los extranjeros se eligieron principalmente con base en el costo, mientras que los locales podían proporcionar entregas más oportunas cuando fuera necesario.

Los esfuerzos cooperativos entre fabricantes y proveedores (así como fabricantes y minoristas en el otro extremo de la cadena) han adquirido cada vez mayor popularidad. Tradicionalmente esta relación fue opuesta. Hoy en día, con la llegada de arreglos como inventarios manejados por el proveedor y MRP II, los proveedores y los fabricantes trabajan muy cerca unos de otros y muchas veces comparten los que alguna vez fueron datos privados para mejorar el desempeño en ambos extremos.

Arreglos para subcontratación

La subcontratación de la fabricación se ha convertido en una popular tendencia en los años recientes. Exitosos fabricantes por contrato como Solectron Corporation experimentaron un rápido crecimiento durante la década pasada. Muchas compañías están subcontratando también funciones de cadena de abastecimiento. La logística de terceros (3PL) se está convirtiendo en un gran negocio. Por ejemplo, Saturn Corporation ahora subcontrata una importante porción de su operación de logística a Ryder Trucks.

Dreifus (1992) describe el caso de una compañía más pequeña que decidió subcontratar las adquisiciones. Singer Furniture Company contrató a la IMX Corporation de Florida para manejar sus compras. En el contrato, IMX maneja todas las negociaciones de embarques consolidados, se encarga del papeleo, conduce inspecciones de control de calidad, y busca proveedores a cambio de una comisión fija. En una corrida de prueba, IMX sugirió un cambio de un proveedor taiwanés de postes de cama a uno localizado en el Caribe, lo que trajo consigo una reducción de 50 por ciento en el precio para Singer. Desde el punto de vista de Singer, las principales ventajas de celebrar este contrato fueron

Economías de escala de compras. IMX compra a varios clientes (incluyendo otros fabricantes de mobiliario), lo que permite negociar varios costos unitarios en pedidos más grandes. Existen también economías de escala en los costos de embarque desde el extranjero. IMX dice que al hacer esto ahorra a sus clientes cuando menos un 20 por ciento.

Menores gastos generales administrativos. Singer pudo eliminar su unidad de compras internacionales y varias funciones administrativas relacionadas.

Menores costos de almacenamiento. IMX corre mucho del riesgo en varias líneas de producto al absorber costos de almacenamiento locales.

Contabilidad en regla. Las firmas estuvieron de acuerdo en abrir procedimientos y desclasificar para asegurar el funcionamiento adecuado de los programas de calidad y que los cargos estuvieran en línea con los costos.

Canales de distribución

El hecho de no establecer canales sólidos de distribución puede conducir al final de muchas compañías. El diseño del canal de distribución incluye el número y configuración de los centros de distribución; arreglos con mayoristas terciarios; otorgamiento de licencias y otros acuerdos con minoristas, incluyendo programas de inventario manejados por proveedores, y el

Aplicación instantánea

DELL COMPUTER DISEÑA LA CADENA DE ABASTECIMIENTO DE VANGUARDIA

Dell Computer fue una de las historias de éxito más grandes de los años noventa. ¿Cómo lo logró? El precio de sus acciones se incrementó un descomunal 27 000 por ciento durante la década. No hay duda de que un diseño de cadena de abastecimiento eficiente fue un factor decisivo a este éxito fenomenal.

Michael Dell, fundador de la compañía, comenzó ensamblando y revendiendo computadoras de escritorio IBM en su habitación de la Universidad de Texas a principios de los años ochenta. Más tarde, a mediados de esa década, formó PC's Limited, que ofreció uno de los primeros "clones" pedidos por correspondencia. Las computadoras eran clones de IBM XT, que había estado vendiéndose en varios miles de dólares en ese momento. PC's Limited vendió una caja básica que consistía en una tarjeta madre con un chip Intel 8088, fuente de energía, unidad de disco flexible y un controlador por $795. El comprador tenía que agregar una tarjeta de gráficos, monitor y unidad de disco duro para que el sistema fuera funcional, pero el costo final del producto era menos de la mitad de su equivalente en IBM XT y corría más rápido. (Esta fue, de hecho, la primera computadora que compró el autor.) Desde sus modestos inicios comenzó a crecer hasta convertirse en la multimillonaria Dell Computer Corporation que domina el mercado de las computadoras hoy en día.

Para entender el éxito de Dell, debemos comprender la naturaleza del mercado de las computadoras personales. El procesador central es realmente la computadora, y su poder determina el poder de la PC. En el mercado de las computadoras, Intel ha sido el líder en diseño de cada nueva generación de chip procesador. Al momento de escribir este texto, Intel está vendiendo la sexta generación de procesadores. La progresión de microprocesadores desde la primera PC de IBM hasta nuestros días es el 8088 (y su similar 8086), 286, 386, 486, Pentium y Pentium II. Algunas compañías, como AMD y Texas Instruments, han tenido un limitado éxito en ese mercado, pero la implacable introducción de nuevos productos por parte de Intel ha dado como resultado que los procesadores de los competidores se vuelvan obsoletos antes de que puedan recoger una importante participación del mercado. Cada nueva generación de microprocesadores vuelve obsoleta la tecnología anterior. Por esa razón, las computadoras y sus componentes tienen un tiempo de vida relativamente corto. Conforme se desarrollan nuevos chips a un paso cada vez más rápido, el problema de la obsolescencia se vuelve aun más crítico.

¿Cómo puede uno evitar quedarse varado con el inventario obsoleto? La solución de Dell fue simple y elegante: ¡no quedarse con inventario! Todas las PC compradas a Dell se hacen bajo pedido. Dell no cuenta con existencias. Sin embargo, sí almacenan componentes. Aunque Dell no puede garantizar que todos los componentes se utilicen, su estrategia de mercado está diseñada para mover tanto como sea posible el inventario de componentes. Se enfoca solamente en la última tecnología, y tanto los sistemas como los componentes reciben un precio adecuado para venderlos rápidamente. Lo que es más, debido a su alto volumen, pueden solicitar descuentos a sus proveedores. Dell se enfoca en usuarios finales prestigiosos que demandan lo último en tecnología y están dispuestos a pagarla. Los márgenes de ganancia son mayores en este mercado, lo que representa a Dell un excepcional desempeño en una industria donde los márgenes son muy estrechos.

Conforme aumenta la llamada velocidad de reloj de la industria (un término acuñado por Fine, 1998), se incrementa también la ventaja de Dell sobre los fabricantes con diseños de abastecimiento más tradicionales. Por ejemplo, Compaq también ha tenido éxito durante la pasada década vendiendo a través de Sears, Circuit City y otros minoristas importantes. Sin embargo, corre el riesgo de tener inventario obsoleto en los canales de distribución de los revendedores. ¿Por qué los fabricantes de PC tradicionales no copian el enfoque de Dell? Como Fine señala, cambiar de un sistema a otro no es fácil. Ese periodo de resaca puede ser fatal si los consumidores experimentan situaciones de escasez de inventario entre sus minoristas locales. (Sin embargo, al momento de escribir este texto, muchos otros fabricantes de computadoras, incluyendo a Compaq y Gateway, ofrecen equipos directamente al consumidor por Internet.)

establecimiento de los medios para mover el producto rápidamente a través de un canal. La venta directa a los consumidores ha sido una estrategia enormemente exitosa para algunos minoristas, especialmente para productos que se vuelven obsoletos muy rápidamente. Al eliminar intermediarios, los fabricantes reducen el riesgo de que el producto muera en el canal de abastecimiento. Esto ha sido una estrategia de mucho éxito para los fabricantes de computadoras como Dell y Gateway y para el vendedor de libros Amazon.com. El fenomenal éxito de Dell Computer en los años noventa puede atribuirse en parte a su excepcional diseño de cadena de abastecimiento, resaltado en la Aplicación instantánea de esta sección.

Problemas para la sección 6.7

18. Describa el concepto de aplazamiento en las cadenas de abastecimiento. Si está planeando viajar a un lugar distante, ¿qué decisiones desearía aplazar lo más posible?
19. Muchos automóviles pueden pedirse con uno de dos tamaños de motor (algunos ejemplos son Lincoln LS, Lexus Coupe y Jaguar S Type) pero son virtualmente idénticos en todos los demás aspectos. ¿Cómo pueden estos fabricantes de autos usar el concepto de aplazamiento en su planeación de producción?
20. Comente por qué tener demasiados proveedores resulta problemático. ¿Puede ser también un problema tener pocos proveedores?
21. Muchas grandes compañías que tienen sus propias instalaciones de fabricación y organizaciones de logística subcontratan una parte de su producción (como IBM) o sus operaciones de cadena de abastecimiento (como Saturn). En su opinión, ¿por qué hacen esto?

6.8 EL PAPEL DE LA INFORMACIÓN EN LAS CADENAS DE ABASTECIMIENTO

¿Cuántas veces hemos escuchado que vivimos en la "era de la información"? La disponibilidad de información está creciendo a una tasa exponencial. Todos los días se introducen nuevas fuentes de información en forma de diarios comerciales y académicos, revistas, boletines informativos, etc. La explosión de disponibilidad de información en la Red ha sido verdaderamente fenomenal. Internet ocupa actualmente el primer lugar en búsqueda de información de casi cualquier cosa para muchas personas.

Conocimiento es poder. En las cadenas de abastecimiento, información es poder. Proporciona a la persona encargada de la toma de decisiones el poder de ir adelante en la competencia, el poder de trabajar en un negocio sin problemas y eficientemente, y el poder de tener éxito en un ambiente cada vez más complejo. La información juega un papel clave en la administración de la cadena de abastecimiento. Como vimos en los capítulos anteriores de este libro, muchos aspectos de la planeación de operaciones comienzan con el pronóstico de ventas y crean un plan para la fabricación o reabastecimiento de inventarios a partir de ese pronóstico. Los pronósticos, desde luego, se basan en la información.

Un excelente ejemplo del papel de la información en las cadenas de abastecimiento son los casos de la Harvard Business School Barilla SpA (A y B) (1994) escritos por Jan Hammond. En muchos cursos de administración de operaciones, los casos proporcionan una introducción al papel de la información en las cadenas de abastecimiento. Barilla es una compañía italiana que se especializa en la producción de pasta. A finales de los ochenta, el encargado de logística de Barilla intentó introducir un nuevo enfoque para tratar con los distribuidores, que denominó distribución justo-a-tiempo (JITD, *just-in-time distribution*). Brevemente, la idea era obtener los datos de ventas directamente de los distribuidores (clientes de Barilla) y usarlos para determinar cuándo y qué tan grandes debían ser las entregas. En ese momento Barilla estaba operando en forma tradicional. Los distribuidores debían colocar de manera independiente los pedidos basándose en los métodos de punto de reorden. Esto llevó a grandes vaivenes en la demanda de las fábricas de Barilla debido al "efecto latigazo", comentado con más detalle a continuación. La idea de JITD encontró mucha resistencia tanto dentro como fuera de Barilla. Sus organizaciones de comercialización y ventas en particular se vieron más amenazadas por los cambios propuestos. Los distribuidores estaban preocupados porque pensaban que sus prerrogativas se verían comprometidas.

Sin entrar en detalles [que podrán encontrarse en Barilla SpA (B)], finalmente prevaleció la administración y el sistema JITD se implementó con varios de los más grandes distribuidores de Barilla. Los resultados fueron sorprendentes. Mejoró la confiabilidad de la entrega y la variación de los pedidos colocados en la fábrica se vio reducida sustancialmente. El programa probó ser una situación de ganar-ganar para Barilla y sus clientes.

El éxito de Barilla con su programa JITD es un ejemplo de lo que hoy en día se conoce como el inventario manejado por el proveedor (VMI, *vendor-managed inventory*). Los programas VMI han constituido el sostén principal de muchos minoristas exitosos en los Estados Unidos. Procter & Gamble ha asumido la responsabilidad de mantener el seguimiento de inventarios de varios de sus principales clientes, como Wal-Mart. Hace algunos años, las administraciones de las tiendas de abarrotes eran responsables de dar seguimiento a sus inventarios. Hoy en día es muy común ver, en nuestras tiendas de abarrotes locales, a personas que verifican los anaqueles sin ser empleados de la tienda sino empleados de los fabricantes. Con los programas VMI se vuelve responsabilidad del fabricante mantener llenos los anaqueles. Aunque los desabastos ciertamente lastiman al minorista, el daño para el fabricante es mayor. Cuando un producto se agota, los clientes normalmente lo sustituirán con otro, así que la tienda realiza la venta de cualquier forma. Es realmente el fabricante quien sufre la penalidad de las ventas perdidas, así que tiene un fuerte incentivo para mantener los estantes con mercancía.

El efecto latigazo

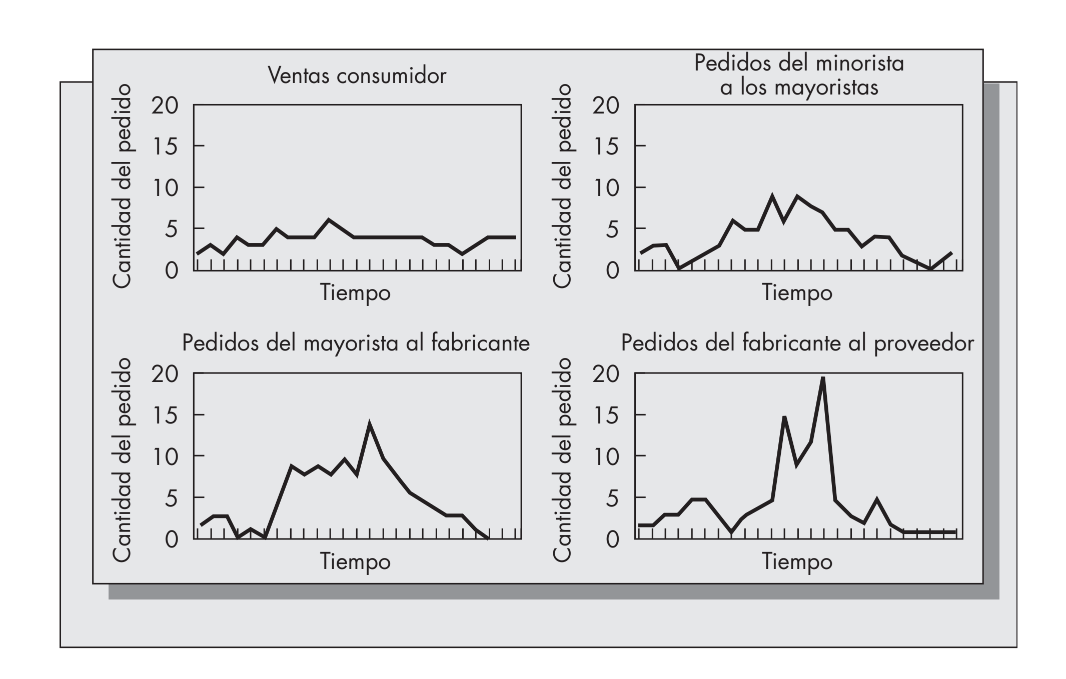
La experiencia de Barilla antes de la implementación de su programa VMI es un ejemplo del efecto latigazo. Se ha convertido en un tema de considerable interés entre los practicantes y los académicos. La historia del latigazo parece ser la siguiente.

Los ejecutivos en Procter & Gamble (P&G) estaban estudiando los patrones de reabastecimiento de sus productos de mayor venta: los pañales desechables Pampers. Se sorprendieron al ver que los pedidos realizados por los distribuidores tenían mucha más variación que las ventas en las tiendas minoristas. Más aún, los pedidos de materiales a los proveedores habían tenido todavía más variabilidad. La demanda de los pañales es muy estable, así que podemos asumir que la variación debería ser menor en toda la cadena de abastecimiento. Sin embargo, éste no era claramente el caso. P&G acuñó el término "efecto latigazo" para este fenómeno. También se le conoce como efecto "trallazo" o de "sierra".

Este fenómeno fue observado también por otras compañías. HP experimentó el efecto latigazo en los patrones de ventas de sus impresoras. Los pedidos hechos por un revendedor mostraron mayores vaivenes que las tiendas minoristas, y los pedidos realizados por la división de impresoras a la división de circuito integrado de la compañía mostraron vaivenes aún más amplios. La figura 6-9 muestra cómo se incrementa la variación conforme nos movemos hacia arriba en la cadena de abastecimiento.

FIGURA 6-9
Variabilidad creciente de los pedidos en la cadena de abastecimiento.

Fuente: H. L. Lee, P. Padmanabhan y S. Whang, 1997.



¿De dónde proviene el efecto latigazo? Una causa surge de un fenómeno que se comentará en el capítulo 7 en los sistemas MRP. Consideremos un sistema básico de dos niveles MRP donde el patrón de la demanda final es fijo y constante. Ya que los artículos se producen en lotes de tamaño Q, el patrón de demanda un nivel abajo es mucho más susceptible. La tendencia de niveles más bajos en la cadena de abastecimiento para procesar pedidos de esta forma es una de las causas centrales del efecto latigazo.

También son comunes en la industria de los abarrotes los altos niveles de inventarios, provocados en parte por el efecto latigazo. Para enfrentar este problema, la industria ha adoptado una iniciativa de respuesta eficiente del consumidor (ECR, *efficient consumer response*) (véase, por ejemplo, Crawford, 1994). La cadena de abastecimiento total de entrega de alimentos, desde el punto en el que los productos dejan al fabricante hasta el punto en que son almacenados en los estantes del minorista, tiene un promedio de 100 días de abastecimiento. La meta establecida de la iniciativa ECR es ahorrar $30 millardos al año al hacer más eficiente la logística de entrega de alimentos.

Otro ejemplo del efecto latigazo es el popular "juego de la cerveza" debido a Sterman (1989). Los participantes juegan los papeles de minoristas, mayoristas y fabricantes de cerveza. La comunicación entre participantes está prohibida: cada jugador debe tomar únicamente decisiones de pedidos sobre lo que le demanda el jugador que se encuentra abajo en la escala. Lo que observamos son importantes vaivenes en los pedidos realizados en el orden descendente, aunque las demandas originales se mantienen prácticamente estables. Esto es consecuencia del efecto latigazo.

La variación telescópica en los patrones de demanda en una cadena de abastecimiento da como resultado una pesadilla de planeación para muchas industrias. ¿Qué puede hacerse para aliviar estos efectos? Primero necesitamos comprender las causas de este fenómeno. Según Lee, Padmanabhan y Whang (1997), las principales causas del efecto latigazo son cuatro:

- Actualización del pronóstico de demanda
- Procesamiento por lotes de pedido
- Fluctuaciones de precios
- Juego de desabasto

Consideraremos cada uno de estos efectos separadamente. Los *pronósticos de demanda* en cada etapa de la cadena de abastecimiento son resultado de las demandas observadas en un nivel inferior (como en el juego de la cerveza). Sólo en la etapa final de la cadena (el minorista) se observan directamente las demandas del consumidor. Cuando cada persona en una cadena de abastecimiento serial pronostica individualmente la demanda genera el efecto latigazo. El minorista construye inventarios de seguridad para protegerse contra la incertidumbre en la demanda. Estos inventarios de seguridad provocan que los pedidos de los mayoristas tengan mayor variación que los de los consumidores. El distribuidor observa estos vaivenes en los pedidos del minorista y crea existencias de seguridad aún más grandes, etcétera.

El *procesamiento por lote* es el fenómeno que vemos en los sistemas MRP y da como resultado patrones de demanda uniformes que se traducen en patrones de demanda susceptibles en niveles más bajos de la estructura de producto. La tendencia natural a ahorrar costos fijos al hacer pedidos con menos frecuencia (la idea detrás de la fórmula CEP) da lugar al procesamiento por lote. Dada la estructura de costos en cada nivel, esta es ciertamente una respuesta razonable.

Cuando los *precios fluctúan* existe un motivo especulativo para mantener inventarios. (Esto se comentó por vez primera en el contexto de administración de inventarios de Arrow, 1958, y se relaciona con las motivaciones para retener efectivo postuladas por el economista Maynard Keynes.) En la industria alimentaria, la mayoría de las transacciones del fabricante a los distribuidores se hacen bajo el arreglo de "compra a futuro". Esto se refiere a la práctica de comprar antes de ser necesario debido a que los fabricantes ofrecen un precio atractivo. Dichas prácticas contribuyen también al efecto latigazo. Se hacen grandes pedidos cuando se ofrecen promociones.

El *juego del desabasto* se presenta cuando no hay abastecimiento de producto y los fabricantes colocan a los clientes en asignación. Cuando los clientes se dan cuenta de que no pueden obtener todas las unidades que desean, sencillamente inflan los pedidos para hacer frente a un déficit anticipado. Por ejemplo, si un fabricante de computadoras espera recibir sólo la mitad de una solicitud para un CPU en poco abastecimiento, simplemente proporcionan el doble del tamaño de su pedido. Si las demandas anticipadas no se materializan, pueden cancelarse los pedidos. El resultado es que el fabricante obtiene una imagen inflada de la demanda real del producto. Esto puede traer consecuencias para el fabricante que invierte grandes cantidades de capital en una expansión de capacidad basada en demandas "fantasma".

Claramente, el efecto latigazo no es resultado de una deficiente planeación o de conducta irracional por parte de los participantes en la cadena de abastecimiento. Cada persona actúa para optimizar su posición. ¿Qué puede, entonces, hacerse para aliviar tal situación? Hay varios remedios potenciales, que comentaremos. Sin embargo, éstos deben tomar en cuenta que las personas se comportan de manera egoísta. Debe eliminarse lo que motiva la conducta que induce el latigazo.

Lee, Padmanabhan y Whang (1997) recomiendan cuatro iniciativas.

1. Intercambio de información
2. Alineación de canales
3. Estabilización de precios
4. Disuasión del juego del desabasto

1. *Intercambio de información* significa que todas las partes involucradas comparten información sobre los datos del punto de venta (POS) y basan sus pronósticos sólo en esos datos. El intercambio de información puede lograrse mediante varias técnicas, una de las cuales es el IED (intercambio electrónico de datos), comentado más adelante con mayor detalle. La tendencia hacia el intercambio de información comienza a aceptarse. Los fabricantes de computadoras, por ejemplo, están requiriendo datos de ventas provenientes de los revendedores (estos datos son acerca del retiro de inventarios de la bodega central). Hasta cierto punto, esperamos ver a los fabricantes relacionados directamente con fuentes de datos POS.

2. La *alineación de canales* es coordinar esfuerzos en la forma de fijación de precios, transportación, planeación de inventarios y propiedad entre los sitios superiores e inferiores en la cadena de abastecimiento. Una tendencia que termina con la alineación de canales es el procesamiento de pedidos en lote. Como observamos antes, los costos fijos motivan los pedidos en lote. Al reducir los costos fijos se obtendrá como resultado pedidos de menor tamaño. Uno de los componentes importantes de los costos fijos es el trabajo administrativo que se requiere para procesar un pedido. Con las nuevas tecnologías, como el IED, estos costos pueden reducirse sustancialmente. Otro factor que motiva grandes lotes es el de las economías a escala de transportación. Desde el punto de vista de unidad, sale más barato hacer pedidos de camión lleno que con media carga. Si los fabricantes permiten a los consumidores hacer pedidos de diversos artículos en una sola carga de camión, se reduce la motivación de hacer pedidos grandes de un solo artículo. Otra tendencia que alienta el pedido de lotes pequeños es la subcontratación de logísticas a terceros. Las compañías de logística pueden consolidar cargas de proveedores múltiples. La subcontratación de logística (como la subcontratación de fabricación) se está expandiendo con gran rapidez.

3. Las promociones de precios motivan a los clientes a comprar grandes lotes y almacenar los artículos para su uso futuro. Como se comentó atrás, es muy común en la industria de abarrotes. Mediante la *estabilización de precios*, los patrones de ventas tendrán menor variación. En la venta al menudeo, el efecto de los precios estables es evidente al comparar los patrones de ventas de minoristas como Macy's, que llevan a cabo promociones frecuentes, y tiendas de almacén como Price/Costco, que ofrecen precios bajos todos los días. Las tiendas bodega experimentan ventas más estables que las tiendas de departamentos. De hecho, para muchas tiendas de departamentos, las ventas promocionales representan la mayor parte de su negocio. Los

Aplicación instantánea

SATURN SURGE COMO LÍDER EN LA INDUSTRIA CON UNA ADMINISTRACIÓN CIENTÍFICA DE LA CADENA DE ABASTECIMIENTO

El éxito de Saturn es toda una historia. A mediados de los años ochenta, GM decidió embarcarse en un emocionante experimento. Junto con los demás fabricantes importantes estadounidenses de autos, desde la década de 1970 había estado perdiendo sin parar participación de mercado ante los japoneses. ¿Cuál fue la clave para el fenomenal éxito de los fabricantes japoneses? Dicho en términos llanos, fueron capaces de cubrir una creciente demanda de automóviles baratos, confiables, eficientes en el consumo de combustible y más pequeños. Para lograrlo, las compañías japonesas (Toyota en particular) desarrollaron un enfoque totalmente nuevo para la fabricación, control de calidad y administración de inventarios. (Véase el comentario del justo-a-tiempo en el capítulo 7, por ejemplo.) GM sintió que tal vez era posible adaptar parte de su pensamiento en su nueva división Saturn, capaz de operar independientemente de la compañía matriz.

Saturn ha tenido muchísimo éxito en todos los rubros. Sus autos rivalizan con la competencia japonesa en precio, calidad y confiabilidad. La satisfacción del cliente está entre las más altas de la industria. Una dimensión importante de la satisfacción del cliente se encuentra en el área de los servicios de reparación. Una medida de satisfacción con los servicios de reparación del distribuidor es la frecuencia con la que los clientes regresan al distribuidor para servicio. Los propietarios de autos Saturn los reparan con los distribuidores con mayor frecuencia que los clientes de otros fabricantes de autos. Es más, la disponibilidad de partes de venta directa de Saturn es la más alta de todos los fabricantes de autos, según una encuesta realizada en 1999.[1]

¿Cómo ha logrado Saturn este desempeño en la industria? Entre otras cosas, diseñó una eficiente cadena de abastecimiento de partes, con base en principios científicos sólidos. El sistema de abastecimiento de partes de Saturn en buena parte se inspira en la clase de principios del justo-a-tiempo. Éste es un sistema basado en la demanda, donde el flujo de partes por la cadena de abastecimiento se dispara por una demanda al nivel del cliente. Las partes se almacenan en un centro de distribución (CD) adyacente a esta fábrica en Tennessee. Los distribuidores de Saturn almacenan un cierto subconjunto de partes, y no todos los distribuidores almacenan la misma mezcla.

Supongamos que un cliente necesita una parte que el distribuidor tiene agotada. El primer lugar al que acude el distribuidor a buscar la parte es entre los miembros de su "grupo de cooperación" local, esto es, un grupo de distribuidores cercanos organizado por la compañía matriz. Si la parte no está disponible dentro del grupo de cooperación, entonces se pide al centro de distribución o directamente al proveedor. Si está disponible en el centro de distribución, puede enviarse al distribuidor en un plazo no mayor de tres días. El sistema de Saturn está diseñado para reaccionar ante la demanda de partes más que a intentar pronosticar la demanda de ellas. Este enfoque tiene sentido en dicho contexto debido a la altamente variable e impredecible demanda de autopartes.

¿Por qué Saturn sobresale en su proceso de entrega de partes? Primero, ha designado coordinadores de flujo de material, personas responsables de coordinar la fabricación y atender las demandas de partes. Segundo, adoptó medidas de desempeño relacionadas directamente con la experiencia del cliente, en lugar de medidas más comunes de la industria. Tercero, la matriz asume la responsabilidad de manejar el inventario de partes con los distribuidores.

El éxito de Saturn es una señal positiva de que las compañías de autos estadounidenses pueden competir con éxito ante las compañías japonesas al alinear sus prioridades y situando al cliente en primer lugar. Quizás esto puede proporcionar una llamada de alerta a la misma GM para repensar sus prioridades y organización.

[1] Cohen, M. A., *et al*. "Saturn's Supply-Chain Innovation: High Value in After Sales Service", *Sloan Management Review*, verano de 2000, pp. 93-101.

principales fabricantes de abarrotes como P&G, Kraft y Pillsbury se encaminan hacia una estrategia precio-valor mientras se alejan de los precios de promoción por la misma causa.

4. Una forma de minimizar los pedidos excesivos como resultado del *juego de desabasto* es asignar conforme a registros de ventas pasadas más que en los pedidos. Esto reducirá la tendencia de los clientes a exagerar en los mismos. Varias compañías (incluida General Motors) se están moviendo en esta dirección.

En resumen, el efecto latigazo es la consecuencia de agentes individuales en la cadena de abastecimiento actuando en pro de sus mejores intereses. Para mitigar los efectos de latigazo,

deben efectuarse varios cambios. Los incentivos deben colocarse en su lugar para aminorar los errores en el pronóstico de demanda, reducir el tamaño de órdenes excesivas en situaciones de asignación y alentar el intercambio de información y la alineación del sistema. Al convertirse estas iniciativas en políticas, todos, especialmente el consumidor, se beneficiarán de los costos reducidos y de una mayor eficiencia en la cadena de abastecimiento.

Comercio electrónico

El comercio electrónico es un término general que designa una amplia gama de métodos para efectuar transacciones sin los sistemas tradicionales basados en papel. Esto incluye IED, correo electrónico, transferencias electrónicas de fondos, publicaciones y tableros electrónicos, bases de datos compartidas y todas las formas de sistemas comerciales basados en Internet (Handfield y Nichols, 1999). El sistema de código de barras de punto de venta hoy común en los supermercados y tiendas minoristas es otro tipo de comercio electrónico. Muchas compañías estadounidenses bien conocidas utilizan actualmente múltiples formas de comercio electrónico. Por ejemplo, Wilder & McGee (1997) reportan que General Electric Corporation, uno de los fabricantes diversificados más importantes del mundo con una amplia variedad de productos, utiliza sistemas de transacciones basados en Internet e IED como parte normal de sus prácticas comerciales en la mayoría de las divisiones. Además de ser usuaria, GE también desarrolla sistemas de cadena de abastecimientos con base electrónica. Su división de servicios de información está trabajando con Netscape para desarrollar un sistema que permita a los usuarios de IED realizar negocios en Internet. No únicamente los monstruos corporativos como General Electric usan Internet para hacer negocios. Muchos de los más pequeños están moviendo sus sistemas de cadena de abastecimiento a cierta forma de comercio electrónico. Como ejemplo, Palevich (1999) informa que Do It Best Corporation, un pequeño distribuidor de productos de ferretería, materiales de construcción y plomería, utiliza IED, sistemas basados en Internet, pronóstico electrónico y sistemas de administración de bodegas para ganar competitividad.

La meteórica subida y caída de la venta minorista por Internet es bien conocida. Durante los últimos años del siglo XX, los IPO de las llamadas compañías puntocom se presentaron casi todos los días y los precios de sus acciones se elevaron de manera constante aunque muchas de estas compañías tuvieron poca actividad. Tal como sucedió con el desastre del bulbo de tulipán en Holanda en la Edad Media, el fondo tenía que caer y así ocurrió. Desaparecieron la mayoría de los vendedores de Internet "puros" (esto es, minoristas de Internet sin los tradicionales almacenes). Uno que sobrevivió y prosperó es Amazon.com. Comenzó como librería de descuento. Al vender desde un solo lugar, puede cosechar los beneficios de la centralización de inventarios (comentada en la sección 6.9) y evitar la costosa inversión de capital en los inmuebles tradicionales. Hoy Amazon ha expandido significativamente su línea de productos y es uno de los minoristas "puros" que han sobrevivido a la quiebra de las puntocom.

El modelo de la venta a menudeo por Internet en esencia es la misma que la venta por catálogos. Los minoristas por catálogo han existido desde hace muchos años, siendo Lands' End y L. L. Bean los de mayor éxito. Sin embargo, es necesario desarrollar, imprimir y enviar por correo con regularidad los catálogos, lo que resulta muy caro, en especial considerando que la mayoría de los catálogos se desechan. Por esta razón, Internet puede tener significativas ventajas sobre los negocios tradicionales de compra por catálogo. Por otro lado, el simple hecho de desarrollar una página web no garantiza que se va a tener éxito. Algunos de los principales portales (AOL y Yahoo, por ejemplo) requieren de grandes cantidades de dinero para permitir el acceso directo a un sitio de comercio-e.

Aunque la venta minorista por Internet ha recibido una gran publicidad en los años recientes, representa apenas una pequeña parte de las actividades que componen la vasta arena del comercio electrónico. Los sistemas de comercio electrónico en Internet entre empresas representa una participación mucho mayor del mercado de comercio electrónico del que tienen los minoristas. Por ejemplo, compañías como Cisco Systems vieron crecer la porción de Internet de sus negocios de 5 por ciento en 1996 a más del 80 por ciento en 2000. Internet tiene el

potencial de cambiar toda la cara de la administración de la cadena de abastecimiento en los próximos años.

Intercambio electrónico de datos

El intercambio electrónico de datos (IED) es una tecnología que permite mejorar las operaciones de la cadena de abastecimiento. IED se refiere a la transmisión de documentos comerciales estándar en un formato predeterminado de la computadora comercial de una compañía a la de sus socios comerciales (Cannon, 1993). Lo que lo separa de otras formas de comercio electrónico es que radica en formatos estandarizados más que en formatos privados. Actualmente los dos estándares son ANSI y EDIFACT.

El IED es una forma en que la compañía puede transmitir información, tal como la de demanda del punto de venta, órdenes de compra e información del estado de inventarios, a los usuarios dentro de la compañía y a los clientes y socios comerciales. Puede servir como medio para aminorar el efecto latigazo antes comentado, ya que proporciona un medio para compartir información entre los participantes en una cadena de abastecimiento. Desde el punto de vista del manejo de la cadena de abastecimiento, su principal ventaja es la velocidad. Al transmitir información deprisa, se reducen los tiempos inactivos. Como se comentó en el capítulo 5, es la incertidumbre de la demanda sobre el tiempo de inactividad de reabastecimiento, lo que da como resultado la necesidad de tener un inventario de seguridad.

Los sistemas IED se han implementado en una amplia variedad de ambientes, pero en general entre las grandes compañías. La razón es que el IED puede ser caro. Requiere un software dedicado y un hardware más avanzado del que poseen muchas compañías pequeñas. Un área en la que el IED se ha implementado son los sistemas de inventario manejados por el proveedor (VMI). Gerber Foods, por ejemplo, utiliza VMI con varios de sus clientes y ha instalado un sistema IED consistente con estos minoristas para asegurar que puede comunicarse con ellos de manera eficiente (Cooke, 1999).

Sistemas de transacciones basados en Internet

Internet proporciona una alternativa atractiva para compañías que consideran invertir en el IED. Internet (conocida también como World Wide Web) es una vasta colección de computadoras y servidores conectados, a los que se puede tener acceso desde cualquier parte del mundo por medio de un portal o salida. Ingresamos a la Red por medio de un navegador, y un dispositivo de búsqueda facilita su uso. Netscape/AOL y Microsoft producen los navegadores más populares, y Yahoo parece ser el buscador más importante. El uso de la Red tanto de empresa a cliente (B2C) como de empresa a empresa (B2B) está creciendo con mucha rapidez. La Red tendrá significativas implicaciones para la administración de la cadena de abastecimiento en los próximos años.

Aunque las compañías B2C han adquirido la mayor notoriedad, es la B2B la que tiene mayor potencial para el futuro. Ésta se refiere a sistemas de transacciones por Internet que permiten a las compañías compartir información electrónica como las cantidades de un pedido, pronósticos y necesidades futuras. Muchas firmas de software ofrecen sistemas basados en Internet. Algunas, como Oracle y SAP, unen el software como parte de su completo almacenamiento de información y sistemas de recuperación. Otras ofrecen productos específicamente diseñados para cadenas de abastecimiento B2B. Tres dignas de mención son i2 Technologies de Dallas, y Agile Software y Ariba, de Silicon Valley. A partir de la información obtenida desde sus páginas de Internet, i2 ofrece productos para distintos sectores de la industria, incluyendo automotores, ventas al menudeo, energía y productos químicos, computadoras y electrónica, metales, productos farmacéuticos, semiconductores y transportación y logística. Agile Software ha diseñado sus productos de forma que aprovechen la ventaja de la tendencia hacia la subcontratación de manufactura. Como Agile cuenta con más fabricantes de equipo original (OEM, *original equipment manufacturers*), alienta a que sus proveedores y facilitadores también utilicen su lista para

mantenerse compatibles. Ariba indica que abastecen a compradores, realizadores de mercado y proveedores.

Cientos de compañías ofrecen soluciones para la cadena de abastecimiento a través de Internet. Muchos adaptan sus productos a segmentos específicos de la industria como Print for Less.com (impresión) y Supply Solution Incorporated (mercado secundario de automotores). Al igual que con cualquier nueva empresa, la mayoría de estas firmas no sobreviven y son pocos los líderes de mercado que surgirán.

¿Cuáles son las ventajas que ofrecen estos sistemas de cadena de abastecimiento? Para empezar, virtualmente cualquier persona tiene acceso a Internet. La velocidad de acceso está creciendo a una tasa estable con la llegada de computadoras y modos más rápidos de transmitir datos. Además, Internet no requiere los complejos protocolos que caracterizan los sistemas IED. Por último, los sistemas basados en la Red pueden montarse e implementarse a una fracción del costo del IED.

Podríamos pensar que los sistemas basados en Internet dominan los sistemas IED debido a estas ventajas, pero este no ha sido el caso hasta ahora. Muchas compañías siguen preocupadas por la seguridad. Los complejos protocolos IED proporcionan una seguridad mayor que la que nos dan los sistemas basados en la Red. Algunas compañías están explorando sistemas híbridos donde intercambian documentos IED por Internet. Esto proporciona a compañías más grandes, con una fuerte inversión en IED, los medios para integrar a los clientes pequeños en sus sistemas existentes (Sliwa, 1999). Sin embargo, debemos creer que con todas las ventajas que ofrecen los sistemas basados en la Red, comenzarán a reemplazar a los sistemas IED en el futuro.

La tecnología RFID proporciona un flujo de producto más rápido

Las etiquetas de identificación de radiofrecuencia (RFID, *radio frequency identification*) son una tecnología emergente que cambiará la forma de transmitir información en una cadena de abastecimiento. Los códigos de barras se han vuelto tan comunes en las ventas de menudeo que suponemos nos han acompañado toda la vida. De hecho, la tecnología de código de barras es relativamente nueva, habiéndose adoptado apenas en 1985. Es probable que las etiquetas RFID reemplacen los códigos de barras en los próximos años.

Las etiquetas RFID se inventaron en 1973, pero no es sino hasta hoy que se vuelven comercialmente viables. Son microchips que funcionan con baterías (etiquetas activas) o señales de radio (etiquetas pasivas). El dispositivo pasivo es menor y más barato y es posible que sea la opción para las aplicaciones de administración de inventarios. Éstos reciben energía de los lectores (entonces "despiertan") y transmiten un código único. Las etiquetas pasivas sólo pueden leerse a distancias cortas, pero proporcionan un medio sencillo de identificación electrónica.

Ya que cuestan aproximadamente $0.50 cada una, su uso resulta demasiado caro a escala masiva. Sin embargo, el costo unitario está disminuyendo rápidamente. Se calcula que cuando alcance los $0.05 o menos, comenzaremos a ver el uso masivo de RFID en todos los sectores de la venta minorista. Cualquier artículo con valor mínimo de $1 o más será etiquetado de esta manera (Granneman, 2003).

Aunque no tan comunes como los códigos de barras (¡todavía!), las etiquetas RFID comienzan a aceptarse en el mercado. Las aplicaciones existentes de tecnología RFID incluyen 1) EZ Pass para pagar cuotas de carreteras o puentes, 2) etiquetado de equipaje en algunos vuelos por Delta Airlines, y 3) etiquetado de contenedores de carga en la mayor parte de los puertos del mundo. Conforme disminuya su costo, veremos un rango mucho más amplio de aplicaciones en el contexto de la administración de la cadena de abastecimiento. La conciliación de los envíos con los conocimientos de embarque o listas de empaque y entrega se puede realizar rápida y exactamente de forma electrónica, con lo que se elimina la necesidad de realizar a mano estas funciones. Al momento de escribir esto, P&G y Wal-Mart estaban experimentando con la tecnología RFID para la identificación de productos de consumo (Dining y Schuster, 2003). (Casualmente, fue Wal-Mart quien encabezó un amplio uso del código de barras en los ochenta, y podría ser de nuevo la que encabezara el uso de la tecnología RFID en ventas al menudeo.)

Desde luego, la tecnología RFID tiene implicaciones mucho más amplias que su aplicación a las cadenas de abastecimiento. Por ejemplo, Solusat, el distribuidor mexicano de VeriChip, un microchip del tamaño de un grano de arroz que se inyecta por debajo de la piel, está comercializando su dispositivo como un medio de identificación de emergencia. El interés en México por este producto es consecuencia de que anualmente desaparecen más de 10 000 niños. El fabricante de Verichip, Applied Digital Solutions, dijo que sus plantas extenderán el servicio Verikid a otros países, incluyendo Estados Unidos (Scheers, 2003). Dichas aplicaciones tienen potencialmente enormes beneficios, pero también pueden usarse de formas que amenazan nuestra privacidad. Será interesante ver hacia dónde nos lleva esta tecnología.

Problemas para la sección 6.8

22. ¿Qué es el efecto latigazo? ¿Cuál es el origen del término?
23. ¿Piensa que la eliminación de intermediarios puede conducir siempre a la eliminación del efecto latigazo? ¿Bajo qué circunstancias puede no funcionar?
24. Comente las cuatro iniciativas recomendadas para aminorar el efecto latigazo en las cadenas de abastecimiento.
25. ¿Cuáles son las ventajas y desventajas de los sistemas de transacciones basados en la Red comparados con el IED?
26. Comente las ventajas y desventajas del menudeo-e comparado con la venta al menudeo tradicional.

6.9 SISTEMAS DE DISTRIBUCIÓN MULTINIVEL

En muchos sistemas de inventarios militares y comerciales de gran tamaño, es común que las existencias se almacenen en diversas ubicaciones. Por ejemplo, el inventario puede fluir del fabricante a las bodegas regionales, de las bodegas regionales a las bodegas locales, y de las bodegas locales a las tiendas o a otras ubicaciones de punto de uso. Determinar la asignación de inventario entre estas ubicaciones múltiples es una importante consideración estratégica, y puede tener un significativo impacto en la base. Un típico sistema de distribución multinivel se muestra en la figura 6-10.

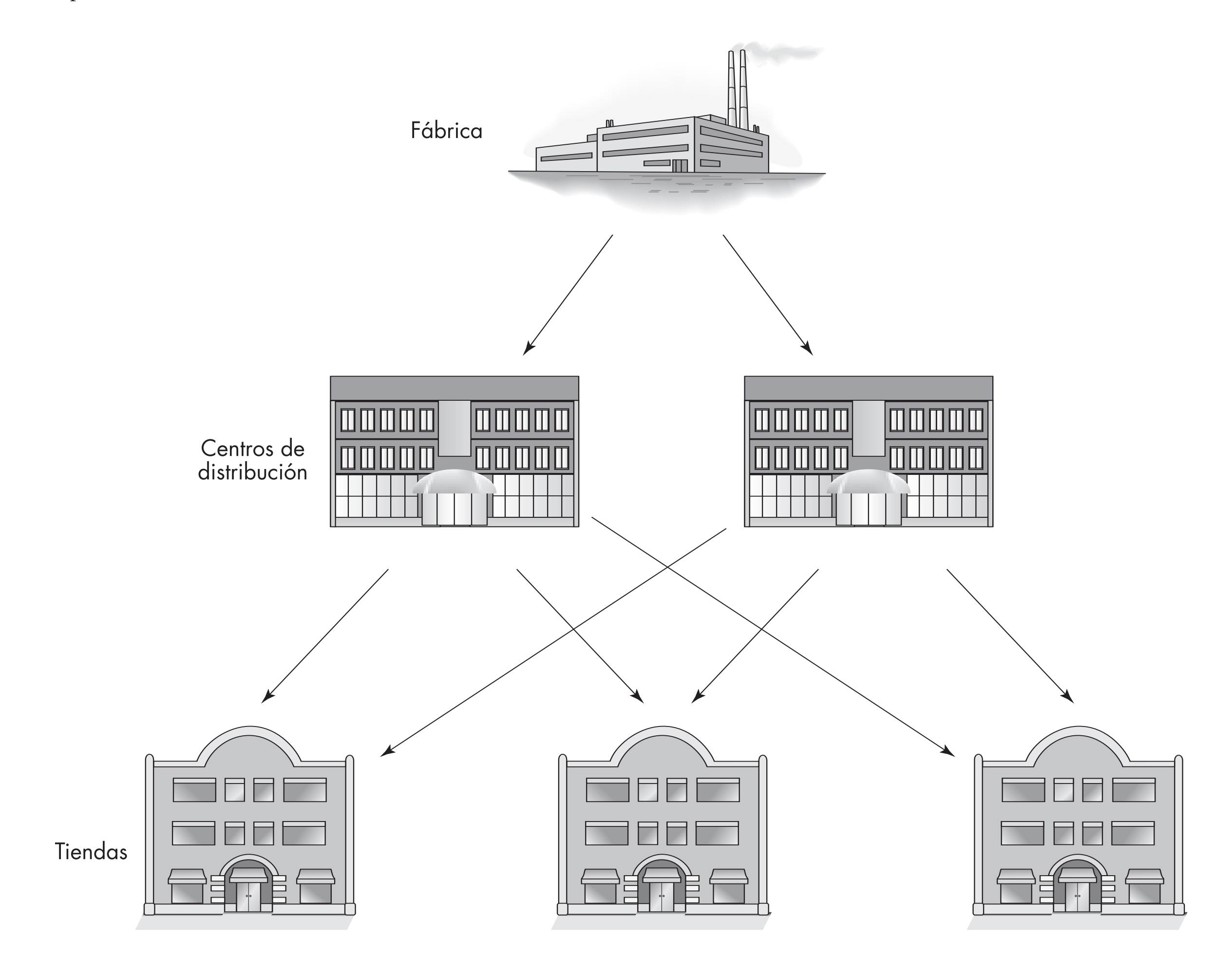
En el lenguaje de la teoría tradicional de inventarios, dichos sistemas se conocen como sistemas de inventario multiescalón. El interés en esta área lo desató la ponencia de Clark y Scarf (1960), parte de una gran iniciativa en Rand Corporation para estudiar los problemas de logística y administración de inventario surgidos en el ejército. Este documento estableció el marco teórico para lo que después sería un amplio cuerpo de investigación en la teoría de inventarios multiescalón. (En el capítulo 5 se presenta un breve panorama general de los modelos de inventario multiescalón.)

El hecho de incluir ubicaciones de almacenaje intermedias en un sistema de distribución de gran escala tiene ventajas y desventajas. Entre las ventajas podemos mencionar

1. Compartición de riesgos.
2. Los centros de distribución pueden diseñarse para cubrir las necesidades locales.
3. Economías de escala en el almacenamiento y movimiento de artículos.
4. Tiempo de respuesta más veloz.

La *compartición de riesgos* se obtiene a partir de que en la mayoría de los casos las demandas agregadas tienen una variación relativa menor que las demandas individuales. (Esta

FIGURA 6-10
Sistema típico de distribución multinivel.



propiedad se comentó también en el contexto del pronóstico en el capítulo 2.) Supongamos que un centro de distribución atiende a 50 puntos de venta minoristas. Al mantener la mayor parte del inventario en el centro de distribución en lugar de en las tiendas, el mismo nivel de servicio de todo el sistema se consigue con menor inventario total. Le mostraremos cómo surge esto con un ejemplo sencillo.

En las ventas al menudeo, la mezcla de productos depende de la ubicación de la tienda. Por ejemplo, se venden más camisas de manga corta en Arizona que en Maine. Por tanto, los *centros de distribución* localizados en el suroeste podrían almacenar una variedad diferente de productos distinta a la de los del noroeste. Los centros de distribución pueden diseñarse para tomar en cuenta estas diferencias.

Los centros de distribución permiten *economías de escala* de la siguiente forma. Si todos los productos se embarcan directamente desde la fábrica al punto de venta local, los tamaños del embarque serían relativamente pequeños. Los embarques voluminosos desde las fábricas localizadas en el extranjero hacia los centros pueden hacerse a centros de distribución a un precio unitario menor. Así, los embarques más pequeños se realizarán en distancias más cortas desde los centros de distribución hacia las tiendas.

Finalmente, debido a que los centros de distribución pueden localizarse más cerca del cliente que las fábricas, las demandas pueden cubrirse con mayor rapidez. En Estados Unidos, los minoristas tratan de localizar los centros de distribución dentro de un día de distancia de

cualquier tienda. (Es por ello que Ohio es un lugar popular para los centros de distribución minorista, ya que es el centro geográfico del país.) De esta forma, los inventarios de las tiendas pueden reabastecerse durante la noche en caso necesario.

Los sistemas de distribución multinivel pueden tener también varias desventajas. Entre las mismas estaría que

1. Pueden requerir de más inventarios que los sistemas de distribución más sencillos.
2. Pueden incrementar los tiempos totales de inactividad del pedido desde la planta hasta el cliente.
3. Pueden finalmente dar como resultado mayores costos para el almacenaje y movimiento de artículos.
4. Pueden contribuir al efecto latigazo.

Según el número de diferentes ubicaciones utilizadas, es probable que el sistema multinivel requiera más inventario total general que un sistema de un solo nivel, ya que es posible incorporar el inventario de seguridad en cada nivel. Esto significa que hay más dinero comprometido en la espera.

Cuando los centros de distribución experimentan desabastos, el tiempo de respuesta del sistema desde el fabricante hasta el punto de venta puede empeorar en un sistema multinivel. Esto es una consecuencia de que, en el caso multinivel, el tiempo total de inactividad de la planta a la tienda es la suma de tiempos inactivos en cada nivel. (Sin embargo, en un sistema bien manejado, los desabastos en el centro de distribución deben ser situaciones muy poco frecuentes.)

No debemos ignorar el papel del centro de distribución en sí. Construir instalaciones modernas a gran escala puede costar hasta $500 millones. Además, hay costos recurrentes como renta y mano de obra. Como resultado, la construcción y mantenimiento de un sistema de distribución multiubicación resultan muy caros.

El *efecto latigazo*, comentado en detalle en la sección 6.8, es la propensión de la variación de pedidos a incrementarse conforme nos movemos hacia arriba en la cadena de abastecimiento. Agregar niveles adicionales a un sistema de distribución puede provocar un efecto latigazo.

Quizá la ventaja más significativa del almacenamiento centralizado es la oportunidad de compartir riesgos, como se comentó antes. Para comprender la forma en que trabaja la compartición de riesgos, consideremos el escenario siguiente. Asumamos que n ubicaciones minoristas independientes almacenan artículos similares. Por ejemplo, éstas pueden ser las tiendas departamentales Macy's localizadas en diferentes ciudades. Asumamos también que el nivel de inventario de un artículo en particular está determinado por el modelo del repartidor de periódicos. Refiriéndonos al comentario de la sección 5.5, existen valores conocidos del costo de excedente de mercancías unitario, c_o, y el costo de escasez de mercancías unitario, c_u. En el capítulo 5 se probó que el nivel óptimo de almacenamiento es la fracción $c_u/(c_u + c_o)$ de la distribución de demanda. (Si suponemos que los niveles de existencias se determinan tomando como base niveles de servicio en lugar de costos, entonces la proporción crítica *es* el nivel de servicio, y c_o y c_u no tienen que conocerse.)

Para simplificar el análisis, supongamos que la demanda de este artículo sigue la misma distribución normal en cada almacén con la media μ y desviación estándar σ, y las demandas son independientes de tienda a tienda. Digamos que z^* es el valor de la variación normal estándar que corresponde a la probabilidad de cola izquierda igual a la proporción crítica. Entonces, como se muestra en la sección 5.3, la política óptima es ordenar hasta $Q = m + \sigma z^*$ en cada ubicación. El inventario de seguridad sostenido en cada ubicación es σz^*, de manera que el inventario de seguridad en el sistema es $n\sigma z^*$. Refirámonos a este caso como el sistema descentralizado.

De manera alternativa, supongamos que todo el inventario para este artículo se mantiene en un solo centro de distribución y se envía a las tiendas por la noche cuando se necesita. Consideremos que la cantidad de inventario de seguridad necesario en este caso para proporcionar el mismo nivel de servicio que en el sistema descentralizado. Ya que las demandas de las tiendas son variables aleatorias normales independientes con media μ y varianza σ^2, la demanda agregada de todas las n tiendas es normal también, pero con la media $n\mu$ y varianza $n\sigma^2$. Por tanto, la desviación estándar de la demanda agregada tiene una desviación estándar $\sigma\sqrt{n}$. Esto

significa que para lograr el mismo nivel de servicio, la bodega necesita almacenarse hasta el nivel Q_w dado por $Q_w = n\mu + z^*\sigma\sqrt{n}$. El inventario total de seguridad es ahora $z^*\sigma\sqrt{n}$. Esto corresponde a un sistema centralizado.

La formación de la proporción del inventario de seguridad en el sistema descentralizado sobre el inventario de seguridad en el sistema centralizado da $z^*\sigma n / z^*\sigma\sqrt{n} = \sqrt{n}$. Por tanto, el sistema descentralizado tendrá que retener $\sqrt{n}$ veces más inventario de seguridad que el sistema centralizado para lograr el mismo nivel de servicio. Incluso para valores menores de n, esta diferencia es significativa; para minoristas grandes como Wal-Mart con miles de tiendas, la diferencia es enorme.

Desde luego este ejemplo es una simplificación de la realidad. Para la mayoría de los productos, no es práctico para cada venta acudir al centro de distribución y no resulta práctico que un solo centro de distribución atienda a todo el país. Más aún, es poco probable que la demanda en diferentes tiendas para el mismo artículo sea independiente. Las demandas de algunos artículos, tales como los artículos de moda deseables, tal vez se correlacionen, mientras que otros, como los artículos estacionales como los trajes de baño, pueden correlacionarse de manera negativa. Sin embargo, pese a estas advertencias, las ventajas de la centralización son sustanciales y ocasionan que los sistemas de distribución multinivel se utilicen ampliamente en muchas industrias, en particular en las de venta al menudeo. Estos resultados se basan en el trabajo de Eppen (1979) (quien admitió las demandas correlacionadas). Este modelo fue ampliado por Eppen y Schrage (1981) y Erkip, Hausman y Nahmias (1990) a ambientes más generales.

Varios autores han examinado la cuestión de cómo las características del artículo afectan el desglose óptimo del centro de distribución contra el inventario en tienda. Muckstadt y Thomas (1980) mostraron que los artículos de alto costo y baja demanda derivaron el beneficio más grande del inventario centralizado, mientras que Nahmias y Smith (1994) mostraron la forma en que otros factores, como la probabilidad de la venta perdida y la frecuencia de los embarques desde el centro de distribución a las tiendas, también afectan el desglose óptimo entre el almacén y el inventario del centro de distribución. Una amplia revisión de los modelos de control de inventario para la venta minorista puede encontrarse en Nahmias y Smith (1993), y una excelente colección de artículos sobre modelos de inventario multiescalón, en Schwarz (1981).

Problemas para la sección 6.9

27. Describa la forma en que la compartición de riesgos trabaja en un sistema de inventarios multiescalón. ¿Bajo qué circunstancias se deriva el mayor beneficio de la compartición de riesgos? ¿Cuál fue el beneficio menor?
28. ¿Por qué un minorista desearía considerar el desarrollo de un sistema multinivel de tres niveles? (Los tres niveles pueden etiquetarse Centro de Distribución Nacional, Centro de Distribución Regional y Tienda.)
29. ¿Cuáles son las características de los artículos desde los cuales derivamos el mayor beneficio del almacenamiento centralizado? ¿Cuál es el beneficio menor?

6.10 DISEÑO DE LA CADENA DE ABASTECIMIENTO EN UN AMBIENTE GLOBAL

Las barreras económicas entre países están desapareciendo en un porcentaje cada vez mayor. Hoy día son pocas las industrias que producen y brindan servicio sólo en el mercado local. La frase "vivimos en una economía global" se ha convertido en un cliché, pero ciertamente es más real que nunca antes. Esto queda demostrado por los hechos siguientes, reportados en Dornier *et al.* (1998):

- Cerca de una quinta parte de la producción de las compañías estadounidenses se realiza en el extranjero.
- Las firmas estadounidenses actualmente tienen $500 millardos en activos de firmas extranjeras. Esta cantidad está creciendo 7 por ciento al año.

- Veinticinco por ciento de las importaciones y exportaciones estadounidenses están entre los afiliados extranjeros y las compañías matrices estadounidenses.
- A principios de los años noventa, las compañías multinacionales representaron 89 por ciento de las ventas de todas las compañías en la base de datos de la Junta de Conferencia.
- Los activos de fabricación de firmas estadounidenses en países extranjeros es cada vez más grande (y viceversa).

Un ejemplo de cambio dramático en el mercado se presentó en la industria automotriz. Considere la experiencia del autor. Cuando era niño en los cincuenta, no había duda de que el auto familiar debía ser estadounidense. De hecho, casi *todo el mundo* en Estados Unidos compraba autos nacionales en ese entonces. Womack *et al.* (1990) informan que a partir de 1955 la participación del mercado nacional estadounidense en manos de compañías estadounidenses era de cerca del cien por ciento. Los fabricantes extranjeros comenzaron a incursionar en el mercado estadounidense a finales de los cincuenta y principios de los sesenta. Fueron principalmente fabricantes europeos (Volkswagen, Mercedes, MG, Volvo, Fiat y Renault), que ofrecieron lo que no ofrecían las firmas estadounidenses. Para 1970, las compañías estadounidenses habían cedido cerca del 10 por ciento del mercado nacional a la competencia extranjera. Durante los setenta, la participación de mercado de las firmas de Estados Unidos se erosionó con mayor rapidez. Mientras que las compañías estadounidenses se aferraron a los viejos diseños y a las antiguas formas de hacer negocios, los fabricantes extranjeros aprovecharon la ventaja del cambio en el gusto. Para 1989, los fabricantes de autos estadounidenses habían perdido una tercera parte del mercado nacional. Los grandes ganadores fueron los japoneses, quienes vieron pasar su participación de mercado de casi cero a mediados de los años cincuenta a cerca del 30 por ciento en 1990 (Womack *et al.*, 1990).

Además del hecho de que los consumidores estadounidenses están comprando más autos extranjeros, también es cierto que las compañías estadounidenses están produciendo más autos en otros países, y los competidores extranjeros están produciendo más autos en Estados Unidos. Una creciente fracción de las plantas ensambladoras estadounidenses se ha reubicado en países con menores costos de mano de obra. GM, por ejemplo, ha invertido fuertemente en plantas de partes en México. Miles de empleados de GM están ganando de 1 a 2 dólares la hora por trabajar en el parque industrial de Matamoros, justo al sur del río Bravo. Al mismo tiempo, los fabricantes extranjeros están construyendo más instalaciones en Estados Unidos. Las plantas japonesas en Estados Unidos producen los dos modelos de mayor popularidad vendidos aquí: el Honda Accord, ensamblado en Marysville, Ohio, y el Toyota Camry, ensamblado en Georgetown, Kentucky. La industria automovilística es una de las muchas donde la producción se ha movido fuera del país y la competencia extranjera ha hecho significativas incursiones en el mercado nacional.

Dornier *et al.* (1998) identifican las cuatro fuerzas conductoras del proceso de globalización:

- Fuerzas del mercado global.
- Fuerzas tecnológicas.
- Fuerzas de costo global.
- Fuerzas políticas y macroeconómicas.

1. Fuerzas del mercado global

Conforme maduran los mercados e ingresan nuevos participantes a los mismos, declinan los márgenes de utilidades. En este punto, las firmas voltean hacia mercados nuevos y menos desarrollados para seguir rentables. Por tanto, un importante componente de las operaciones de la cadena de abastecimiento global es establecer nuevos mercados en el extranjero. Existen varias cuestiones que deben considerarse al hacer esto. Primero, ¿el producto puede venderse "como es", o se necesita adaptarlo para que encaje en el mercado? Algunos productos necesitan muy poca o nula adaptación local. Coca-Cola es un excelente ejemplo. Otros productos deben adaptarse para cumplir con los requerimientos locales. Por ejemplo, los productos electrónicos

deben cumplir con el voltaje e interfaz electrónica. Finalmente, existen productos donde no se requiere de ninguna adaptación local pero la misma es buen negocio. Por ejemplo, tanto GM como Ford venden muchos autos en Europa. Estos modelos están diseñados para satisfacer los gustos europeos y muchas veces son significativamente distintos de los que se venden a nivel nacional. Debido a que los precios de la gasolina son mucho más altos en Europa que en Estados Unidos y muchas ciudades continúan teniendo calles muy estrechas, los europeos prefieren autos compactos y con motores pequeños.

2. Fuerzas tecnológicas

No hay duda de que los avances tecnológicos han contribuido a la globalización. El rápido crecimiento de la tecnología de la información aminora varias desventajas de la ubicación de operaciones en el extranjero. Como se comentó en el capítulo 1, llegar pronto al mercado es una ventaja competitiva muy importante. Localizar la producción y las actividades de investigación y desarrollo en diferentes lugares puede resultar problemático. Una comunicación cercana entre el personal de investigación y desarrollo y el de fabricación es clave para lograr un desarrollo exitoso del producto. Hoy en día, sin embargo, las distancias físicas ya no son un problema. La comunicación electrónica es ahora posible en cualquier parte del mundo por medio del correo electrónico, fax, teléfono y los sistemas de comunicación basados en la Red.

Los avances en la fabricación y logística también han contribuido a la globalización. Generar productos diferentes en distintos lugares para segmentos de mercado variados se hace más sencillo gracias a los avances en la tecnología de fabricación. Las plantas pueden dividirse en módulos y construirse con relativa rapidez en otros países. La variedad de producto puede incrementarse.

La difusión en nuevas tecnologías también contribuye a la globalización. Las firmas pueden buscar ubicar las instalaciones de fabricación cerca de fuentes de tecnologías críticas. Por ejemplo, tanto IBM como Xerox decidieron ubicar sus instalaciones de fabricación de monitores en Japón, el líder en tecnología de monitores. Muchas compañías europeas y japonesas tienen operaciones en Silicon Valley para estar cerca de la sede de las nuevas tecnologías que se desarrollan ahí.

Compartir la nueva tecnología es un factor que motiva alianzas estratégicas entre compañías de diferentes países. Por ejemplo, Read-Rite Corporation de Milpitas, California, firmó una alianza estratégica con Sumitomo Steel de Japón. Read-Rite es uno de los productores independientes de cabezales de disco Winchester, y Sumitomo estaba interesado en adquirir parte de dicha tecnología e irrumpir en tal mercado. A cambio, Read-Rite recibió entrada al mercado japonés y una cantidad sustancial de efectivo (Yansouni, 1999).

La mayor parte de las inversiones estadounidenses en el extranjero son instalaciones de fabricación, como las plantas de General Motors en México, diseñadas para aprovechar la ventaja de la mano de obra barata. Sin embargo, cada vez es más común que las compañías estadounidenses ubiquen también en el extranjero sus instalaciones de investigación y desarrollo. Esto está motivado principalmente por la disponibilidad de personal de bajo costo y altamente capacitado en cuestiones técnicas en países como Taiwán.

3. Fuerzas de costo global

Otro factor importante para que una compañía decida globalizarse es que se eleven los costos de la fabricación nacional. Al comprimirse las utilidades, las compañías buscan no solamente expandir sus mercados sino reducir los costos unitarios. Esta es la principal razón por la que los países desarrollados ubican sus instalaciones en países menos desarrollados. Sin embargo, en años recientes, el proceso parece haberse acelerado mucho más. Ello se debe a muchos factores. Uno es la cada vez mayor competencia global de los competidores internacionales. Otra es el crecimiento de industrias de alta tecnología que construyen instalaciones en el extranjero como parte de su modelo inicial de negocios. Un tercer factor es el mejoramiento de las condiciones en el extranjero. Un ejemplo es que el nivel educativo promedio ha mejorado de manera constante en los años recientes en países como Malasia y Singapur. También, las oportunidades

Aplicación instantánea

TIMKEN HACE FRENTE A LAS IMPORTACIONES CON LA OFERTA POR LOTES

Los fabricantes estadounidenses se ven amenazados por la competencia extranjera. Incluso los fabricantes de autos estadounidenses admiten con facilidad que están comprando muchas de sus partes a proveedores europeos y japoneses. (Ejemplo de ellos: Jaguar, ahora propiedad de Ford, compra productos electrónicos de Japón y transmisiones de Alemania.) Una estrategia que ayuda a los fabricantes estadounidenses a competir es la *oferta por lotes*. Un caso es Timken Company de Randleman, Carolina del Norte. Produce cojinetes cilíndricos grandes, que se usan para reducir la fricción en camiones de volteo y molinos de viento. En años recientes, Timken ha estado expuesta a la competencia proveniente de Japón y Europa Oriental, que vende productos comparables a un precio unitario mucho menor. Timken ha pedido a la administración Bush que vigile las prácticas de *dumping* por parte de estos competidores. Mientras tanto, la compañía no se ha quedado esperando.

Normalmente, los cojinetes y el equipo de lubricación que los acompaña se venden como partes separadas. Por lo general se comprarán a proveedores separados, y el comprador será responsable de unirlos en el producto final. Para evitar el problema al cliente, Timken une los cojinetes con el mecanismo de lubricación y los vende como un sistema integrado. En otras aplicaciones, los fabricantes están tomando partes estándar y las rodean con revestimientos, clavijas, dispositivos de lubricación y sensores electrónicos. Llevando la oferta en lote un paso más allá, algunos fabricantes brindan mantenimiento y servicios de ingeniería con él, proporcionando así servicios y productos que los competidores extranjeros no pueden. Aunque el costo unitario de las partes que comprenden un lote puede ser menor en el extranjero, el valor agregado del lote y los servicios con él, compensa más que de sobra las diferencias en los precios.

Desde el punto de vista del consumidor, la oferta por lotes reduce los gastos generales y mejora la calidad del producto. Según Daniel Murphy, de Caterpillar, Inc., "la oferta en lote estimula la calidad sustancialmente. Es una fuente de reducción de costo y produce una mejor confiabilidad". La oferta por lotes también aminora el número de proveedores.

Existe sólida evidencia de que la estrategia de oferta por lotes está rindiendo frutos. Las compañías que venden sistemas integrados a los fabricantes de autos obtuvieron mejores resultados que aquellas que fabrican productos de consumo, conforme a un estudio realizado a más de 100 proveedores de autopartes por la compañía Oracle y la Oficina para el Estudio de la Transportación Automotriz de la Universidad de Michigan. El ofrecimiento en lote es un ejemplo de la clase de estrategias innovadoras que están ayudando a la industria estadounidense con las importaciones de bajo costo.

Fuente: Tejada, C. "The Allure of 'Bundling'", en *The Wall Street Journal*, octubre 7, 2003.

de nuevos mercados y de mano de obra barata en países como China y la ex Unión Soviética están disponibles desde hace algunos años.

Sin embargo, Dornier *et al.* (1998) dicen que esta estrategia puede ser errónea. Se basa, comentan, en prácticas contables inapropiadas que inflan la importancia de las tasas de mano de obra directa. Conforme cambian las tasas de mano de obra, dicha estrategia lleva a saltar de un lugar a otro para buscar las tasas más bajas. Sin embargo, hacer esto requiere de sustanciales desembolsos de capital para nuevas instalaciones, y existen costos de arranque en la nueva ubicación.

El tratamiento fiscal preferencial es también un incentivo para ubicarse en el extranjero. La mayoría de las compañías farmacéuticas estadounidenses tienen sus principales instalaciones de producción en Puerto Rico, pues el gobierno de Estados Unidos otorga una situación fiscal favorable a las compañías estadounidenses que lleven a cabo ahí sus negocios. Irlanda es otro país que ha proporcionado sustanciales incentivos fiscales para que las compañías se ubiquen allí.

4. Fuerzas políticas y macroeconómicas

Los principales convenios comerciales han ayudado a derribar las barreras de los acuerdos internacionales. Los dos más importantes son el TLC en Norteamérica y la CEE en Europa. Los oponentes políticos al TLC, que con frecuencia hablan del "gigante chupando salud", aludiendo a la pérdida de trabajos nacionales, han probado estar equivocados. Medir los efectos directos

de los acuerdos comerciales es difícil debido a factores confusos, pero las tasas de desempleo ciertamente no se dispararon como resultado del TLC. El desarrollo de la Comunidad Económica Europea ha hecho mucho más sencillo para los fabricantes entablar negocios con Europa, gracias a la estandarización y reducción de aranceles intranaciones.

Las políticas gubernamentales juegan un importante papel en la atracción de una región o de un país específico. China, en particular, proporciona sustanciales incentivos para aquellos que vienen de fuera y ofrecen asociarse con organizaciones nacionales, incluyendo empresas propiedad del gobierno, privadas y colectivas. Otros países que ofrecen sustanciales incentivos para la inversión extranjera son Irlanda, Malasia, Irlanda, México y Tailandia. Sin embargo, los disturbios políticos han contribuido al lento desarrollo de la inversión extranjera en África. La inestabilidad política puede ser un riesgo sustancial al invertir en el extranjero.

Uno de los impedimentos para realizar negocios en el extranjero es el riesgo del tipo de cambio. Si una fábrica construye en China debe pagar los materiales comprados localmente y la mano de obra en yuanes. Si vendemos en Alemania, los recibos están en euros. Conforme el valor relativo de las monedas varía en el mercado mundial, tanto los costos como los recibos de ventas también cambian en relación con la moneda del país. Estos cambios pueden ser impresionantes. Por ejemplo, hasta los años setenta, los dólares canadiense y estadounidense estaban prácticamente a la par. Hoy en día, el tipo de cambio aproximadamente es de $1.35 dólares canadienses por $1 dólar estadounidense. Una inversión de capital en fábricas instaladas en Canadá perdió una tercera parte de su valor en dólares estadounidenses. Desde luego, esto ocurrió en un periodo de tiempo muy largo. Pero las fluctuaciones en el tipo de cambio y las devaluaciones de la moneda pueden, y así ocurre, presentarse mucho más rápido. Proporcionan un riesgo adicional para las compañías que hacen negocios en el extranjero.

Administración de la cadena de abastecimiento en un entorno global

La discusión hasta ahora ha considerado las motivaciones y problemas que surgen de las operaciones internacionales en general. Muchas de estas cuestiones aplican también en la administración de la cadena de abastecimiento global. Por ejemplo, el tipo de cambio juega un importante papel en el diseño de la cadena de abastecimiento. Aunque son un riesgo, las relativas valuaciones de la moneda y otros factores financieros como las tasas de interés locales también pueden verse como una oportunidad. Las compañías pueden "jugar" el sistema de mantener las instalaciones de manufactura en varios lugares y producir donde los costos sean menores.

Las distancias geográficas también presentan problemas especiales para la administración de la cadena de abastecimiento. Aunque hemos aprendido a mover la información a la velocidad de la luz, el tiempo requerido para mover un artículo de un lugar a otro no ha cambiado significativamente en las últimas décadas. Ubicar instalaciones en el extranjero significa mayores tiempos de inactividad para el abastecimiento. Los requerimientos de procedimiento, retrasos burocráticos y condiciones climáticas son factores que incrementan la incertidumbre de los tiempos de inactividad.

Los sistemas de producción como JIT se vuelven mucho más difíciles de implementar cuando las instalaciones y los proveedores se encuentran lejos unos de otros. Una de las razones de la exitosa implementación de JIT en las plantas de Toyota en la ciudad de Toyoda fue que los proveedores se localizaban muy cerca de la planta. Imagine las dificultades al tratar de implementar una disposición JIT con un proveedor que se encuentra a 10 mil kilómetros de distancia.

Las inadecuaciones de infraestructura también presentan problemas especiales para manejar las cadenas de abastecimiento globales. La lista incluye la falta de trabajadores capacitados, malas carreteras, escasez de aeropuertos adecuados, y condiciones de vida indeseables para la administración extranjera. Los sistemas inadecuados de comunicaciones así como los sistemas de telefonía que no cumplen con los estándares y la falta de disponibilidad de una red contribuyen a las dificultades. Dichos factores no sólo ponen freno a la ubicación de instalaciones de fabricación en los países en desarrollo, sino que también entorpecen el establecimiento de canales de distribución.

Aplicación instantánea

DIGITAL EQUIPMENT CORPORATION USA UN MODELADO MATEMÁTICO PARA PLANEAR SU CADENA DE ABASTECIMIENTO GLOBAL

Digital Equipment Corporation (DEC) era una de las compañías de computadoras verticalmente integradas más grandes del mundo. DEC creció rápidamente en los años setenta y se convirtió en uno de los proveedores líderes de las minicomputadoras. Durante la década de 1990 DEC continuó siendo un importante participante en el área de las minicomputadoras, pero vio erosionarse su participación de mercado debido al desplazamiento de que fueron objeto las minicomputadoras a manos de las estaciones de trabajo y las PC. Al final de los noventa, Compaq Corporation adquirió DEC. Antes de esto, DEC enfrentó el problema de manejar una compleja cadena de abastecimiento a escala mundial. Más de la mitad de sus ingresos provienen de 81 países fuera de Estados Unidos lo que la obligó, como a muchas otras compañías internacionales, a manejar de manera eficiente una cadena de abastecimiento global en todo el planeta. Para ayudar con este complejo proceso de planeación, la compañía formó un equipo para considerar la cadena de abastecimiento global de DEC. Algunos de los aspectos que debían solucionarse en la planeación de DEC eran la ubicación de clientes y proveedores, la ubicación y disponibilidad de mano de obra barata y capacitada, la duración de la tramitación de material, las consideraciones fiscales, regulaciones fiscales y los requerimientos de contenido local. El equipo de investigación desarrolló una descripción matemática de la cadena de abastecimiento de DEC en forma de un programa lineal entero-mixto. (La programación lineal se analiza en este texto en el suplemento 1. El problema de transporte, estudiado en este capítulo, es un ejemplo de los tipos de modelos desarrollados por esta aplicación.)

El modelo matemático, llamado GSCM, se desarrolló originalmente para planear la producción y distribución de nuevos productos. Los autores informan que para la primavera de 1994 se había implementado para cerca de 20 productos nuevos. Esto se hizo por medio de un plan de 18 meses para reestructurar el proceso de fabricación y distribución a fin de reducir costos, inventarios y mejorar el servicio al cliente. Como resultado de este proceso, DEC redujo el número de plantas de 33 a 12 mientras mantenía una corriente de ingresos y producción cada vez mayor. Para 1995, como resultado de la implementación de este enfoque, DEC experimentó una disminución en los costos anuales de fabricación de $167 millones y en el costo anual de logística de más de $200 millones.

Fuente: Arntzen *et al.*, 1995.

Nota: Desde 1995 DEC fue adquirida por Compaq, que más tarde se fusionó con Hewlett-Packard.

Como se comentó anteriormente en este capítulo, el diseño puede verse afectado por consideraciones de la cadena de abastecimiento global. La diferenciación retrasada es un medio para mejorar el desempeño de un sistema en el que difieren los requerimientos locales. Hewlett-Packard ha tenido mucho éxito vendiendo impresoras en el extranjero de esta manera. También se exagera el efecto latigazo por los largos tiempos de demora experimentados en las cadenas de abastecimiento global. La operación a escala internacional presenta muchos desafíos para la administración de las cadenas de abastecimiento.

Problemas para la sección 6.10

30. En su opinión, ¿cuáles serán los nuevos mercados emergentes en el mundo en los próximos 20 años? ¿Por qué razón África se ha desarrollado con tanta lentitud como nuevo mercado y como una ubicación deseable para la fabricación?
31. Según se comentó con anterioridad en este capítulo, la participación en ventas en Estados Unidos de las firmas multinacionales se está incrementando. ¿Qué eventos pueden revertir esta tendencia?
32. La fabricación ha continuado yéndose al extranjero con mucha rapidez. ¿Por qué algunos escritores sienten que ésta es una estrategia inadecuada? ¿Qué factores pueden revertir esta tendencia?
33. ¿Qué dificultades han surgido para la administración de la cadena de abastecimiento en virtud del crecimiento de la globalización?

6.11 Resumen

¿Por qué ha surgido interés en la administración de la cadena de abastecimiento en los años recientes? Como dice James Alampi, presidente de Van Waters and Rogers, el distribuidor de productos químicos más grande de Estados Unidos:

> Francamente, el costo de hacer un producto es casi irrelevante. Se tiene mayor oportunidad de recuperar el costo en la cadena de abastecimiento de la que se tiene por la fabricación. Existe mucha duplicación e ineficiencia. (Henkoff, 1994)

El señor Alampi hace una observación que muchos otros están comenzando a contemplar. Aunque la fabricación continúa siendo un problema importante, la administración está notando que un mejor diseño de la cadena de abastecimiento es donde pueden realizarse sustanciales ahorros en los costos. Debido a esto, el software dedicado a las cadenas de abastecimiento se está convirtiendo en un gran negocio. Por ejemplo, i2 Technologies de Texas produce e instala software dedicado para cadenas de abastecimiento. Al momento de escribir esto, la compañía está creciendo a más del 50 por ciento anual y cuenta entre sus clientes a firmas como GE, IBM, Dell, TI, y muchas otras. i2 no es la única compañía de software que aprovecha la ventaja del enorme interés en la administración de la cadena de abastecimiento. Algunos de los nombres más importantes en la arena de los sistemas de información, incluidos Oracle Corporation y la alemana SAP, ahora ofrecen módulos de cadena de abastecimiento como parte de su sistema total de soluciones.

El modelado matemático puede jugar un importante papel en la administración eficiente de la cadena de abastecimiento. Los problemas de transporte y trasbordo, estudiados en este capítulo, son ejemplo de la clase de modelos de optimización matemática que pueden ayudar a las grandes compañías a determinar calendarios eficientes para mover el producto de la fábrica al mercado. Existen también técnicas basadas en las matemáticas para una programación eficiente de los vehículos de entrega. El método de ahorros, comentado en este capítulo, es un sencillo ejemplo de una técnica de ese tipo.

La entrega del producto al consumidor solía ser una consideración secundaria en las compañías de manufactura. Hoy en día, sin embargo, las consideraciones de la cadena de abastecimiento se toman en cuenta inclusive a nivel de diseño de producto. Los productos que pueden apilarse con facilidad se pueden almacenar y embarcar con mayor sencillez y menor costo. La estrategia de aplazamiento ha resultado ser un principio fundamental de diseño probado como altamente efectivo en costo. Al diseñar productos cuya configuración final puede aplazarse, las compañías pueden retrasar la diferenciación del producto. Esto permite ganar un tiempo muy valioso para determinar la realidad de la demanda, y ayuda a evitar el problema de producir una mezcla de producto que no se equipare a las ventas.

Se ha observado un fenómeno inusual a finales de los años ochenta, denominado efecto latigazo. La variación de los pedidos parece incrementarse espectacularmente conforme se avanza en la cadena de abastecimiento. Aunque la mayor parte está de acuerdo en que éste es resultado de diferentes agentes que actúan para optimizar sus propias posiciones, las soluciones a este problema no se entienden muy bien. Proporcionar los incentivos a los diversos participantes para que compartan información y reduzcan el procesamiento del pedido puede ser una forma de combatir el efecto latigazo, pero continúa en cuestión si estos métodos funcionan realmente en la práctica.

Una importante tendencia es el movimiento hacia soluciones de la cadena de abastecimiento que aprovechan los enormes pasos dados en la industria de la información. Éstos incluyen tecnologías como el IED y los sistemas de transacciones basados en la Red. Los sistemas de entrada de punto de venta basados en el código de barras son ahora de uso común en la mayor parte de las grandes operaciones minoristas.

Bibliografía

Arntzen, B. C., G. G. Brown, T. P. Harrison y L. L. Trafton. "Global Supply Chain Management at Digital Equipment Corporation". *Interfaces* 25, núm. 1 (enero-febrero 1995), pp. 69-93.

Arrow, K. J. "Historical Background". Capítulo 1 de *Studies in the Mathematical Theory of Inventory and Production*, K. J. Arrow., S. Karlin y H. Scarf, comps. Stanford, CA: Stanford University Press, 1958.

Barrone, F. Comunicación personal, 1996.

Bell, W. J. "Improving the Distribution of Industrial Gases within an Online Computerized Routing and Scheduling Optimizer". *Interfaces* 13, núm. 6, 1983, pp. 4-23.

Cannon, E. *EDI Guide. A Step by Step Approach.* Nueva York: Van Nostrand Reinhold, 1993.

Clark, A. J. y H. E. Scarf. "Optimal Policies for a Multiechelon Inventory Problem". En *Management Science* 6, 1960, pp. 475-490.

Clarke, G. y G. W. Wright. "Scheduling of Vehicles from a Central Depot to a Number of Delivery Points". *Operations Research* 12, 1964, pp. 568-581.

Cohen, M. P. V., Kamesam, P. Kleindorfer, H. Lee y A. Tekerian. "Optimizer: IBM's Multi-Echelon Inventory System for Managing Service Logistics". *Interfaces* 20, núm. 1, enero-febrero de 1990, pp. 65-82.

Cooke, J. A. "Software Takes Babel out of Vendor Managed Inventory". *Logistics Management & Distribution Report* 38, núm. 2, febrero de 1999, p. 87.

Copacino, W. C. *Supply Chain Management: The Basics and Beyond.* Boca Raton, FL: St. Lucie Press, 1997.

Crawford, F. A. "ECR: A Mandate for Food Manufacturers?" *Food Processing*, febrero de 1994.

Dinning, M. y E. W, Schuster. "Fighting Friction". *APICS—The Performance Advantage*, febrero de 2003.

Dornier, P-P., R. Ernst, M. Fender y P. Kouvelis. *Global Operations and Logistics*, *Text and* Cases. Nueva York: John Wiley & Sons, 1998.

Dreifus, S. B., comp. *Business International's Global Desk Reference.* Nueva York: McGraw-Hill, 1992.

Duff, C. y R. Ortega. "How Wal-Mart Outdid a Once-Touted Kmart in Discount Store Race". *The Wall Street Journal*, marzo 24, 1995.

Eilon, S. C. D. T., Watson-Gandy y N. Christofides. *Distribution Management: Mathematical Modeling and Practical Analysis*. Londres: Griffin, 1971.

Eppen, G. D. "Effects of Centralization on Expected Costs in a Multi-Location Newsboy Problem". *Management Science* 25 (1979), pp. 498-501.

Eppen, G. D. y L. Schrage: "Centralized Ordering Policies in a Multi-Warehouse System with Leadtimes and Random Demand". *Multi-Level Production/Inventory Systems: Theory and Practice*, L. B. Schwarz, comp. Nueva York: North Holland, 1981.

Erkip, N., W. H. Hausman y S. Nahmias. "Optimal Centralized Ordering Policies in Multi-Echelon Inventory Systems with Correlated Demands". *Management Science* 36 (1990), pp. 381-392.

Fine, C. H. *Clockspeed.* Reading, MA: Perseus Books, 1998.

Granneman, S. "RFID Chips Are Here", *The Register* (www.theregister.co,uk), consultado el 27 de junio de 2003.

Hammond, Jan. Barilla SpA (A and B). Copyright © 1994 por el presidente y asociados de Harvard Business School.

Handfield, R. B. y E. L. Nichols Jr. *Introduction to Supply Chain Management*. Upper Saddle River, NJ: Prentice Hall, 1999.

Henkoff, R. "Delivering the Goods". *Fortune*, noviembre 25, 1994, pp. 64-78.

Jacobson, G. y J. Hillkirk. *Xerox, American Samurai.* Nueva York: Macmillan, 1986.

Lee, H. L., C. Billington y B. Carter. "Hewlett-Packard Gains Control of Inventory and Service through Design for Localization". *Interfaces* 23, núm. 4 (julio-agosto 1994), pp. 1-11.

Lee, H. L., P. Padmanabhan y S. Whang. "The Bullwhip Effect in Supply Chains". *Sloan Management Review*, primavera de 1997, pp. 93-102.

Martin, A. J. *DRP: Distribution Resource Planning*, 2a. ed., Essex Junction, VT: Oliver Wight Limited Publications, 1990.

Muckstadt, J. A. y L. J. Thomas. "Are Multi-Echelon Inventory Methods Worth Implementing in Systems with Low Demand Rate Items?" *Management Science* 26, 1980, pp. 483-494.

Nahmias, S. y S. A. Smith. "Mathematical Models of Retailer Inventory Systems: A Review". *Perspectives in Operations Management*, R. K. Sarin, comp. Boston: Kluwer, 1993.

Nahmias, S. y S. A. Smith. "Optimizing Inventory Levels in a Two-Echelon Retailer System with Partial Lost Sales". *Management Science* 40, 1994, pp. 582-596.

Palevich, R. F. "Supply Chain Management". *Hospital Material Management Quarterly* 20, núm. 3, febrero 1999, pp. 54-63.

Ragsdale, C. T. *Spreadsheet Modeling and Decision Analysis*, 2a. ed. Cincinnati: South-Western, 1998.

Ross, David Frederick. *Competing through Supply Chain Management: Creating Market-Winning Strategies through Supply Chain Partnerships.* Nueva York: Chapman & Hall, 1998.

Shrage, L. "Formulation and Structure of More Complex/Realistic Routing and Scheduling Problems". *Networks* 11, 1981, pp. 229-232.

Schwarz, L. B., comp. *Multi-Level Production/Inventory Systems: Theory and Practice.* Nueva York: North Holland, 1981.

Scheeres, J. "Tracking Junior with a Microchip", *Wired News* (www.wired.com), consultado el 10 de octubre de 2003.

Signorelli, S. y H. Heskett. "Benneton". Harvard Business School, Caso 9-685-014, 1984.

Simchi-Levi, D., P. Kaminski y E. Simchi-Levi. *Designing and Managing the Supply Chain: Concepts, Strategies, and Case Studies*. Nueva York: McGraw-Hill/Irwin, 1999.

Sliwa, C. "Users Cling to EDI for Critical Transactions". *Computerworld* 33, núm. 11 (marzo 15, 1999), p. 48.

Sterman, R. "Modeling Managerial Behavior: Misperception of Feedback in a Dynamic Decision Making Experiment". *Management Science* 35, núm. 3, 1989, pp. 321-339.

Supply Chain Forum. http://www.stanford.edu/group/scforum/, consultado en agosto de 1999.

Wilder, C. y M. K. McGee. "GE: The Net Pays Off". *Informationweek*, febrero de 1997, pp. 14-16.

Womack, J. P., D. T. Jones y D. Ross. *The Machine That Changed the World.* Nueva York: Harper Perennial, 1990.

Yansouni, C. Comunicación personal, 1999.

Capítulo **Siete**

Sistemas de producción de empuje y tracción: MRP y JIT

Perspectiva general del capítulo

Objetivo

Entender las filosofías de empuje y tracción en la planeación de la producción y comparar los métodos MRP y JIT para programar el flujo de bienes en una fábrica.

Puntos clave

1. *Empuje contra tracción.* Existen dos filosofías fundamentales para desplazar material en la fábrica. Un sistema de empuje es aquel donde la planeación de la producción se realiza por adelantado en todos los niveles. Una vez terminada la producción, las unidades se promueven al siguiente nivel. Un sistema de tracción es aquel donde los artículos se desplazan de un nivel al siguiente sólo cuando se requiere. La *planeación de requerimiento de materiales* (MRP, *materials requirements planning*) constituye el sistema básico de empuje. El sistema de planeación MRP determina cantidades de producción para cada nivel del sistema, sobre la base de predicciones relacionadas con artículos finales en cierto horizonte de planeación específico. Este sistema se basa en el denominado *cálculo de la explosión*, que requiere el conocimiento del factor gozinto ("va hacia", es decir, qué cantidad de la parte A se requiere para la parte B), y en los tiempos de entrega. El primero de los sistemas de tracción es *kanban*, creado por Toyota, que dio como resultado los movimientos *justo a tiempo* (JIT) y la producción esbelta. En este caso, el objetivo fundamental consiste en reducir los productos en proceso a un nivel mínimo. Para lograrlo, los artículos solamente se desplazan cuando lo requiere el siguiente nivel superior en el proceso de producción. Cada uno de los métodos tiene ventajas y desventajas.
2. *Fundamentos de MRP.* El cálculo de la explosión MRP es un conjunto de reglas para convertir un plan *maestro de producción* (MPS, *master production schedule*) en uno diseñado para todos los componentes, incluyendo el producto final. El MPS es un plan de producción para el artículo final o producto final por periodo. Se deriva de los pronósticos de demanda ajustada para utilidades, inventarios disponibles e información semejante. En cada etapa del proceso, se calcula la producción requerida en cada nivel del proceso de producción efectuando dos operaciones básicas: 1) compensar el momento en que inicia la producción con el tiempo de entrega que se requiere en el nivel actual; 2) multiplicar la demanda del nivel superior por el factor gozinto. El plan más simple de producción en cada nivel es lote por lote (L4L, *lot-for-lot*), lo cual significa que se produce la cantidad de unidades que se requieren en cada periodo. Sin embargo, si se conocen los costos de almacenamiento y preparación de la producción, es posible construir un plan de tamaño

de lote más eficiente. Las tres heurísticas que analizamos son: 1) la determinación del tamaño de lote por CEP; 2) la heurística de Silver-Meal, y 3) la heurística de costo unitario mínimo. El ajuste del tamaño de lote óptimo requiere una programación dinámica y se analiza en el apéndice 7-A. También consideramos el ajuste del tamaño del lote cuando las restricciones de capacidad se toman en cuenta explícitamente. Este problema resulta difícil de resolver de forma óptima, pero es posible llegar a aproximarse adecuadamente.

El MRP posee ventajas y desventajas como sistema de planeación sobre otros. Entre sus desventajas están que 1) no se toma en cuenta la incertidumbre del pronóstico; 2) las restricciones de capacidad se pasan completamente por alto; 3) elegir el horizonte de planeación puede tener un efecto significativo en los tamaños de lote recomendados; 4) los tiempos de entrega se consideran fijos, aunque deberían depender de los tamaños de lote; 5) no toma en cuenta las pérdidas debidas a artículos defectuosos o tiempo de inactividad de máquina; 6) la integridad de los datos puede ser un problema serio; 7) en los sistemas donde los componentes se utilizan en diversos productos, es necesario vincular cada orden a un artículo de nivel superior.

3. *Fundamentos de JIT.* La filosofía JIT surgió del sistema kanban creado por Toyota. *Kanban* es una palabra japonesa que significa tarjeta o boleto. Kanban controla el flujo de bienes en la planta empleando una diversidad de tarjetas. Cada tarjeta se pega a una paleta de bienes. La producción no puede dar inicio hasta que los kanban ordenados para la producción se encuentren disponibles. Esto garantiza el hecho de que la producción en cierto nivel no dé inicio a menos que exista demanda en el siguiente. Esto evita que los inventarios de productos en proceso se acumulen entre los centros de trabajo cuando se presenta un problema en el sistema. Parte de lo que hizo que kanban tuviera tanto éxito en Toyota fue el desarrollo del cambio de dados en un solo minuto (SMED, por sus siglas en inglés), que redujo los tiempos de cambio en ciertas operaciones de horas a minutos. Kanban no es la única forma de configurar un sistema JIT. Los flujos de información se pueden controlar mejor con un procesador central de información que con tarjetas.

4. *Comparación de JIT con MRP.* JIT posee ventajas y desventajas cuando se compara con MRP como sistema de planeación de la producción. Entre las ventajas de JIT están que 1) reduce los inventarios de productos en proceso, lo cual reduce los costos de inventario y desperdicio; 2) es fácil identificar rápidamente problemas de calidad antes de que se generen inventarios grandes o partes defectuosas; 3) cuando se coordina con un programa de compra JIT, garantiza el libre flujo de materiales en la totalidad de proceso de producción. Las ventajas de MRP incluyen las siguientes: 1) la capacidad de reaccionar a los cambios en la demanda, ya que los pronósticos de demanda constituyen parte integral del sistema (en contraste con JIT, que no incluye planeación a futuro); 2) holgura para el tamaño del lote en diversos niveles del sistema, dando oportunidad de reducir preparaciones y sus costos, y 3) planear los niveles de producción a todos los niveles de la empresa para varios periodos futuros, lo cual brinda a la compañía la oportunidad de prevenir con el fin de planear mejor los cambios de turno y ajustar los niveles de la fuerza de trabajo ante la demanda cambiante.

La cadena de abastecimiento es el conjunto de actividades que convierten la materia prima en el producto final. Una de las actividades clave en la cadena de abastecimiento es el proceso de producción en sí. La forma en que se dirige la fábrica desempeña un papel fundamental en la confiabilidad y calidad del producto final. Existen dos filosofías relativas al manejo del flujo de bienes en la fábrica, que son fundamentalmente diferentes. Como veremos en este capítulo, los métodos de los capítulos 4 y 5 concernientes al manejo de inventarios no siempre resultan adecuados en el contexto de la fábrica.

Los dos enfoques que consideramos son la *planeación de requerimientos de materiales* (MRP) y *justo a tiempo* (JIT). A menudo estos reciben el nombre de sistemas de control *de*

empuje y *de tracción,* respectivamente. Con el fin de apreciar exactamente lo que distingue a los sistemas de empuje y tracción, necesitamos entender exactamente cómo funcionan estos métodos, los cuales estudiaremos minuciosamente en este capítulo. La definición más simple que el autor ha visto (la cual se debe a Karmarkar, 1989) es la siguiente: "un sistema *de tracción* da inicio a la producción como reacción a la demanda presente, mientras que un sistema *de empuje* da inicio a la producción como anticipación a la demanda futura". De esta manera, MRP incorpora pronósticos de la demanda futura, mientras que JIT no lo hace.

Para comprender mejor la diferencia entre MRP y JIT, consideremos el siguiente ejemplo. Una planta en Muncie, Indiana, fabrica azadones para jardín. Cada azadón consta de dos partes: el cavador metálico y el mango de madera. Las partes se unen por medio de dos tornillos. La planta produce azadones a una velocidad promedio de 100 por semana. El cavador metálico se fabrica en grupos de 400 en los primeros dos días de cada mes, y los mangos se solicitan a un proveedor externo. El ensamble de los azadones se realiza durante la primera semana de cada mes.

Considérese ahora el esquema de demanda de los tornillos. Se necesitan exactamente 800 tornillos durante la primera semana da cada mes. Suponiendo que un mes tiene cuatro semanas, el esquema de demanda semanal para los tornillos es de 800, 0, 0, 0, 800, 0, 0, 0, 800, 0, 0, 0, y así. Supongamos que la fórmula CEP arroja una cantidad de pedido de 1 400, con una velocidad de demanda semanal promedio de 200 y costos adecuados de almacenamiento e instalación. Un análisis simple muestra que ordenar los tornillos en lotes de 1 400 no tiene mucho sentido. Si programamos una entrega de 1 400 tornillos a principios de mes, 800 se utilizan inmediatamente y 600 se almacenan para uso posterior. Al principio del siguiente mes, se ordenan otros 1 400, ya que los 600 tornillos almacenados son insuficientes para satisfacer la demanda del siguiente mes. Esto tiene más sentido para cualquier orden de 800 tornillos a principios de cada mes o algún múltiplo de 800 cada cierta cantidad de meses.

La solución CEP resultó evidentemente inadecuada en este caso. ¿Por qué? Recordemos que al deducir la fórmula CEP, supusimos que la demanda era conocida y constante. En este caso el esquema de demanda se conoce, aunque evidentemente no es constante. De hecho, es muy complicado. Si aplicáramos los métodos del capítulo 5, supondríamos que la demanda es aleatoria. Es fácil demostrar que, a lo largo de un periodo de un año, la demanda semanal posee una media de 200 y una desviación estándar de 350. Estos valores podrían utilizarse para generar valores (Q, R), suponiendo alguna forma de distribución para la demanda semanal. No obstante, esta solución no tendría ningún sentido. El esquema de demanda para los tornillos no es aleatorio; es predecible, ya que resulta del plan de producción para los azadones, que se conoce. La demanda es *variable,* pero no *aleatoria.*

Aún no hemos resuelto el problema sobre la cantidad de tornillos que se necesita comprar, ni cuándo deben entregarse. Una posibilidad consistiría en sencillamente ordenar a principios del año con el fin de satisfacer la demanda para los 12 meses. Esto implicaría una entrega única de 10 400 a principios de cada año (suponiendo 200 semanales). ¿Cuál sería la ventaja de este enfoque? Los tornillos son artículos poco caros. Si se compran suficientes para una producción anual, se incurriría en gastos fijos de entrega una sola vez al año.

Existe una forma completamente diferente de abordar el problema. Podríamos simplemente decidir programar entregas de tornillos a principios de cada mes. Este enfoque podría ser más costoso que la estrategia de entrega una vez al año, porque los costos fijos serían 12 veces más altos. Sin embargo, podría tener otras ventajas, que compensarían la elevación de los costos fijos. Las entregas mensuales eliminan la necesidad de almacenar tornillos en la planta. Si las tarifas de servicio varían, el tamaño de las entregas podría ajustarse a las necesidades. Asimismo, si surge un problema con los tornillos por causa de algún defecto en la producción o un cambio en el diseño de los azadones, la compañía no se vería afectada por un gran inventario de artículos inútiles.

Estas dos políticas ilustran las diferencias básicas entre MRP y JIT (aunque, como veremos, estas filosofías de control de producción abarcan mucho más que esto). En un sistema

MRP, determinamos tamaños de lote sobre la base de pronósticos de demandas futuras y, posiblemente, sobre la base de consideraciones de costos. En un sistema JIT, intentamos reducir los tamaños de lote al mínimo con el fin de eliminar desperdicios y la acumulación innecesaria de artículos.

El MRP puede considerarse un sistema de planeación vertical donde las decisiones sobre las cantidades de producción se deducen de pronósticos de demanda. Las decisiones sobre tamaños de lotes se encuentran en cada nivel del sistema de producción. Los artículos se producen sobre la base de este plan y se promueven, *empujan*, al siguiente nivel. En el JIT, las solicitudes de bienes tienen origen en un nivel superior del sistema y se *jalan*, a través de varios niveles de producción. Ésta es la idea fundamental detrás de los sistemas de empuje y tracción para el control de la producción.

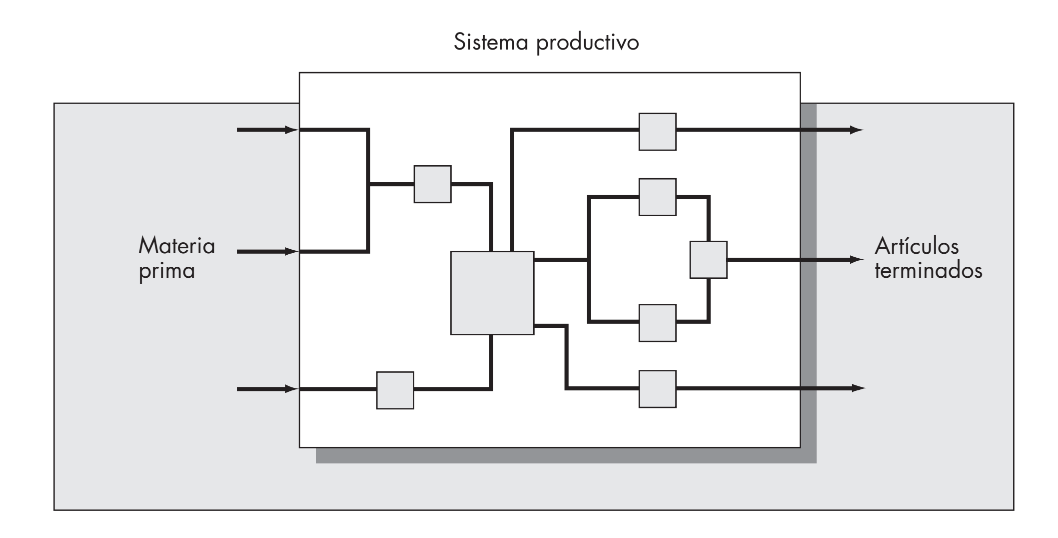
Fundamentos de MRP

En general, un plan de producción es una especificación completa de la cantidad de artículos terminados o productos finales y producidos por subconjuntos, de los tiempos exactos, de los tamaños de lote de producción, y del programa final de terminación. El plan de producción puede descomponerse en varias partes: 1) el **programa maestro de producción (MPS),** 2) el **sistema de planeación de requerimientos de materiales (MRP), y** 3) el **programa detallado de trabajos en el piso de producción.** Cada una de estas partes puede representar un subsistema grande y complejo dentro del plan general.

En el corazón del plan de producción se encuentran los pronósticos de demanda de los artículos finales producidos durante el horizonte de planeación. Un producto terminado constituye la salida del sistema de producción; es decir, los productos enviados. Los componentes son artículos en etapas intermedias de la producción y la materia prima consta de los recursos que entran al sistema. La figura 7-1 muestra un esquema del sistema de producción. Es importante recordar que la materia prima, los componentes y los productos terminados se definen en un sentido relativo y no absoluto. De ahí que pretendamos aislar una parte de la operación de la compañía como un sistema de producción. Los productos terminados asociados con una parte de la compañía pueden ser materia prima para otra porción de la empresa. Un sistema productivo único puede ser una operación de manufactura de la compañía o sólo una pequeña parte de ésta.

El programa maestro de producción (MPS) especifica las cantidades exactas y los tiempos de producción de cada artículo terminado en un sistema productivo. El MPS se refiere a artículos *no acumulados.* De esta manera, las entradas para determinar el MPS son pronósticos de la demanda futura por artículo, más que de los artículos acumulados, como se describió en el capítulo 3. Así, el MPS se descompone en un programa detallado de producción para cada componente que abarca un artículo terminado. El sistema de planeación de requerimientos de manufactura (MRP, *materials requirements planning*) es el medio por el cual se consigue esto.

FIGURA 7-1
Esquema del sistema productivo.



Finalmente, los resultados del MRP se convierten en programas específicos para el piso de producción (empleando métodos como los del capítulo 8) y requerimientos para la materia prima.

Las fuentes de datos para determinar el MPS incluyen lo siguiente:

1. Pedidos en firme de los clientes.
2. Pronósticos de demanda futura por artículo.
3. Requerimientos de inventario de seguridad.
4. Planes estacionales.
5. Pedidos internos de otras partes de la organización.

Una parte importante del éxito del MRP es la integridad y oportunidad de los datos. El sistema de información que sustenta al MRP recibe apoyo de los departamentos de producción, mercadotecnia y finanzas de la compañía. El flujo sin problemas de la información entre estas tres áreas en funcionamiento constituye el ingrediente fundamental para que un sistema de planeación de la producción tenga éxito.

Podemos considerar que el control del sistema de producción consta de tres fases principales. La fase 1 consiste en reunir y coordinar la información que se requiere para formular el programa maestro de producción. La fase 2 es determinar las liberaciones de los pedidos planeados utilizando el MRP; la fase 3 reside en crear un programa detallado de trabajos del piso de producción y de requerimientos de recursos de acuerdo con las autorizaciones de pedidos planeadas con MRP. La figura 7-2 muestra un esquema de estas tres fases de control del sistema productivo.

El capítulo tiene que ver con la forma en que se utiliza el MPS como entrada al sistema MRP. Mostraremos detalladamente cómo funciona el cálculo del MRP; es decir, la forma en que las estructuras de producción se convierten en relaciones de padre a hijo entre niveles del sistema de producción, la forma en que los tiempos de entrega se utilizan para obtener requerimientos en fases del tiempo, y la forma en que los métodos para determinar tamaños de lote dan como resultado programas específicos. En la sección sobre la determinación de tamaños de lote consideraremos las técnicas de determinación de tamaños de lote óptima y la heurística en el caso de sistemas no capacitados, y una técnica heurística directa para la determinación capacitada del tamaño de lote.

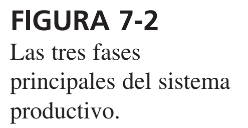
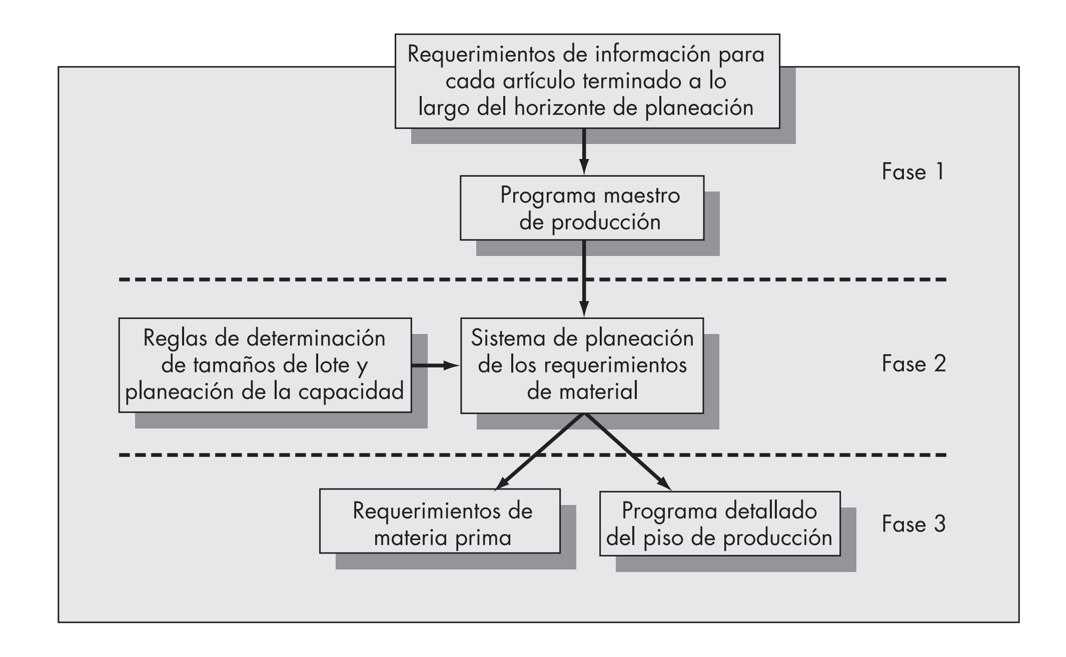


FIGURA 7-2
Las tres fases principales del sistema productivo.



Fundamentos de JIT

El enfoque justo a tiempo tiene sus raíces en el sistema kanban de flujo de materiales iniciado por Toyota. En la sección 7.6 analizaremos el funcionamiento de kanban. La noción de JIT se ha ampliado considerablemente desde su origen como tecnología de flujo de materiales. Produce efectos estratégicos importantes para las firmas, no solamente de manufactura, sino en cuanto a la dirección de la base de proveedores y la administración de la distribución.

Las ideas fundamentales detrás de JIT son:

1. El *inventario de productos en proceso se reduce a un mínimo.* La cantidad permitida de este inventario constituye una medida del grado de afinación del sistema JIT. Mientras menos inventario de productos en proceso se considere en el sistema, mejor balanceadas tendrán que estar las diversas etapas del proceso.
2. *JIT es un sistema de tracción*. La producción en cada etapa se inicia sólo cuando se requiere. El flujo de información en un sistema JIT se transfiere de un nivel a otro de forma secuencial.
3. *JIT se extiende más allá de los límites de la planta.* Las relaciones especiales con los proveedores tienen que estar en orden con el fin de garantizar que las entregas sean realizadas según se necesiten. Proveedores y fabricantes deben localizarse en las proximidades, si el diseño JIT ha de incluir a los proveedores.
4. *Los beneficios de JIT van más allá del ahorro de los costos por inventario.* Las plantas pueden funcionar de manera eficaz sin acumular existencias de materia prima y bienes terminados parcialmente que atasquen el sistema. Los problemas de calidad pueden identificarse antes de que adquieran proporciones incontrolables. La reelaboración y la inspección de bienes terminados se reducen al mínimo.
5. *El enfoque JIT requiere un compromiso serio de la dirección, así como de los trabajadores.* Los trabajadores necesitan mantenerse al tanto de sus sistemas y productos, además de contar con el poder de parar el flujo de la producción si observan alguna anomalía. La dirección debe autorizar a estos trabajadores dicha flexibilidad.

Últimamente se ha empleado el término *producción esbelta* para describir el sistema JIT. El término parece haber sido acuñado por Womack y otros (1990) en su sobresaliente estudio sobre la industria automotriz, *The Machine That Changed the World* (La máquina que transformó el mundo). Al comparar lo peor de la producción en masa norteamericana con lo mejor de la producción esbelta japonesa, los autores muestran cuán eficaz puede resultar una filosofía JIT adecuadamente planteada. Así narran su experiencia en la planta Framingham (Massachusetts) de General Motors en 1986:

> Ahora echemos un vistazo a la línea misma. Cerca de cada estación de trabajo había pilas —en algunos casos el equivalente de semanas— de existencias. Tiradas por doquier había cajas desechadas y otro material provisional para envolver. En la línea misma el trabajo se encontraba distribuido irregularmente: algunos trabajadores corrían como locos para no interrumpir el trabajo y otros hallaban tiempo para fumar e incluso leer el periódico… Al final de la línea encontramos lo que parece la mejor evidencia de anticuada producción en masa: una enorme área de trabajo repleta de automóviles terminados llenos de defectos. Todos estos automóviles requerían más arreglos antes de enviarse, una tarea que requiere enorme cantidad de tiempo y, con frecuencia, implica fallas en la reparación de los problemas, sepultados bajo capas de partes y tapicería.

Ahora comparemos esto con la experiencia en la planta Takaoka de Toyota en Toyoda City:

> Las diferencias entre Takaoka y Framingham impresionan a cualquiera que comprenda la lógica de la producción esbelta. Para comenzar, difícilmente se veía a alguien en los pasillos. Los ejércitos de trabajadores indirectos tan visibles en GM no aparecían por ninguna parte, y, prácticamente, cada trabajador que estaba a la vista se encontraba mejorando el automóvil… La línea de ensamble final revelaba más diferencias. Cerca de cada trabajador en Takaoka se encontraba menos del equivalente de una hora de existencias. Las partes continuaban su viaje sin complicaciones y las tareas estaban mejor equilibradas, de modo que cada trabajador laboraba aproximadamente al mismo paso… Al final de la

línea, la diferencia entre la producción esbelta y la producción en masa resultaba todavía más sorprendente. En Takaoka casi no observamos ningún área de reprocesado. Casi cada automóvil se transfería directamente de la línea al barco o a los camiones que entregan los automóviles al comprador.

Estas diferencias resultan emocionantes y sorprendentes. Debemos hacer notar que desde entonces GM mantiene cerrada la planta de Framingham y que las plantas de las *tres grandes* (a saber, GM, Ford y la división Chrysler de Daimler-Chrysler) son mucho más eficientes y se encuentran mejor dirigidas que Framingham. Aun así, todavía contamos con una forma de duplicar el éxito fenomenal de los japoneses con los sistemas de producción esbelta.

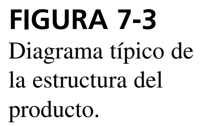
Este capítulo inicia con un análisis sobre el cálculo básico de la explosión del MRP y la forma en que las estrategias de ajuste del tamaño de lote por lote se incorporan en una solución básica MRP de un solo nivel.

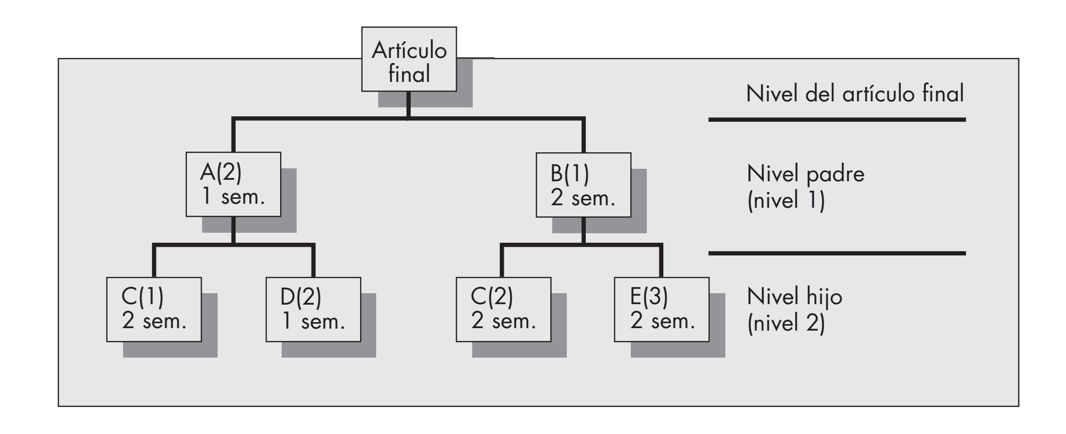
7.1 CÁLCULO DE LA EXPLOSIÓN

Cálculo de la explosión es un término que se refiere al conjunto de reglas por medio de las cuales la mayor parte de los requerimientos en un nivel de la estructura del producto se traducen en un programa de producción para ese nivel y en requerimientos en niveles inferiores. En el corazón de cualquier sistema MRP se encuentra la estructura del producto. La estructura del producto es la relación entre los componentes en niveles adyacentes del sistema. El diagrama de estructura del producto detalla la relación padre-hijo entre los componentes y los artículos en cada nivel, la cantidad de periodos que se requieren para la producción de cada componente y el número de componentes que se requieren en el nivel hijo para producir una unidad en el nivel padre.

Una estructura de producto típica aparece en la figura 7-3. Con el fin de producir una unidad del artículo final, se requieren dos unidades de A y una unidad de B. El ensamble de A requiere una semana, y el ensamble de B, dos semanas. A y B son *hijos* del artículo final. Para producir A se necesitan una unidad de C y dos unidades de D. Producir B exige dos unidades de C y tres unidades de E. Los tiempos respectivos de demora también aparecen en el diagrama de estructura del producto. Los diagramas de estructura del producto pueden resultar complejos y tener hasta 15 o más niveles en algunas industrias.

El cálculo de la explosión (también conocido como explosión de la lista de materiales) sigue un conjunto de reglas que convierten las requisiciones planeadas de artículos finales y componentes en programas de producción de componentes de nivel inferior. El método en sí implica el "faseo" o colocación progresiva en el tiempo de los requerimientos y explicar la cantidad de componentes que se necesitan en el nivel hijo con el fin de producir un solo artículo padre. El método se ilustra mejor con un ejemplo.


FIGURA 7-3
Diagrama típico de la estructura del producto.



Ejemplo ***7.1***

Harmon Music Company produce una variedad de instrumentos de viento en su planta en Joliet, Illinois. Como la compañía es relativamente pequeña, quisiera reducir al mínimo la cantidad de dinero que gasta en inventario. Por esa razón se establecen niveles de producción que se ajustan lo más posible a la demanda predicha. Con el fin de conseguir este objetivo, la compañía adopta un sistema MRP para determinar cantidades de producción.

Uno de los instrumentos producidos es la trompeta modelo 85C. El costo de menudeo de la trompeta es de $800 y ha resultado un artículo razonable si no espectacularmente rentable para la compañía. Sobre la base de pedidos de las tiendas de música en el país, el gerente de producción recibe pronósticos respecto a la futura demanda de alrededor de cuatro meses.

La figura 7-4 muestra la trompeta y sus diversos componentes. La figura 7-5 proporciona el diagrama de estructura del producto de la construcción de la trompeta. La sección de campana, el tubo principal y las secciones con válvulas se unen en el ensamble final. Antes de la unión, los tres ensambles corredizos y las tres válvulas se fabrican y se ajustan al conjunto de la caja de las válvulas. El moldeado e impresión de forma de la sección de campana requiere dos semanas, y el moldeado e impresión de forma del tubo principal y las secciones de válvulas requieren cuatro semanas. Las válvulas requieren tres semanas de fabricación, y los conjuntos corredizos dos semanas.

El problema del ensamble de la trompeta es un sistema MRP de tres niveles. El nivel 0 corresponde al producto terminado o artículo final, que es la trompeta completa. El nivel 1, nivel hijo en relación con la trompeta, corresponde a los ensambles de la campana y la caja de las válvulas. El nivel 2 corresponde a los ensambles corredizos y de válvulas. La información del diagrama de estructura del producto a

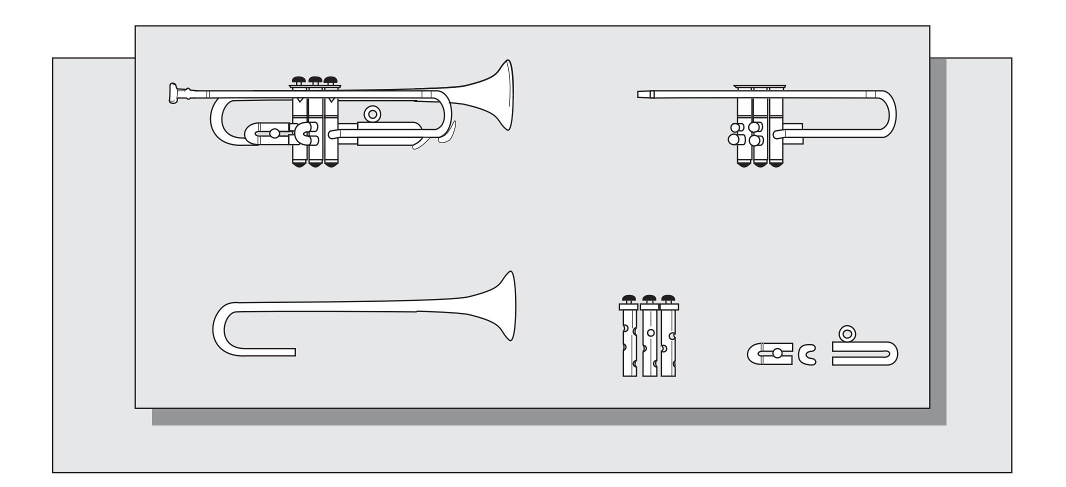


FIGURA 7-4
Trompeta y sus subensambles.

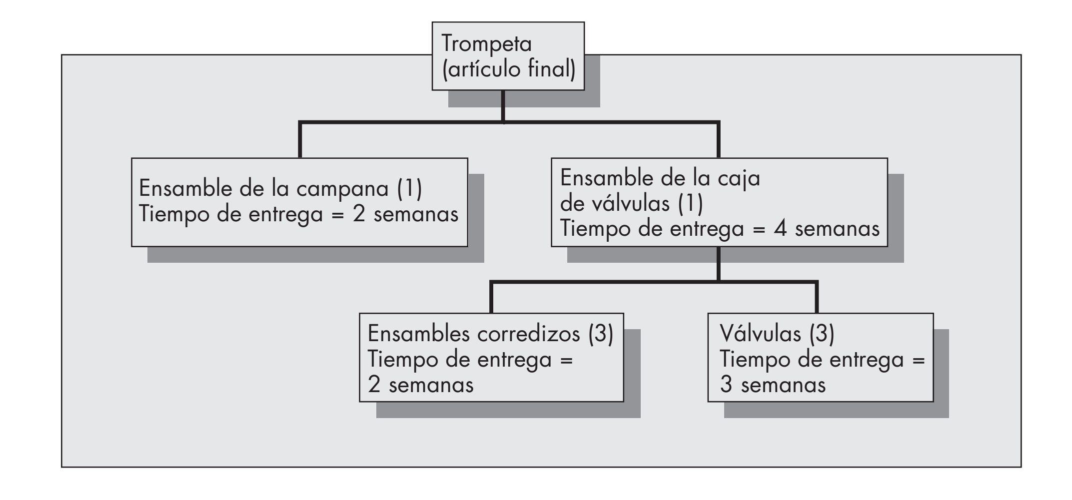


FIGURA 7-5
Diagrama de estructura del producto para una trompeta Harmon.

menudo se presenta como una lista indexada de materiales (BOM, *bill of materials*), que constituye una representación más conveniente para preparar la reintroducción a una computadora. La lista de materiales indexada para la trompeta es[1]:

1 trompeta
  1 ensamble de la campana
  1 ensamble de las válvulas
    3 ensambles corredizos
    3 válvulas

Se requieren siete semanas para producir una trompeta. De ahí que la compañía deba iniciar ahora la producción de trompetas que se enviarán en siete semanas. Por esa razón sólo consideraremos pronósticos de demandas a siete semanas. Si designamos la semana actual como la semana 1, entonces Harmon requiere pronósticos de ventas de trompetas para las semanas 8 a 17. Supongamos que las demandas predichas para tales semanas son:

Semana	8	9	10	11	12	13	14	15	16	17
Demanda	77	42	38	21	26	112	45	14	76	38

Estos pronósticos representan las cantidades de trompetas que la compañía quisiera tener listas para enviar en las semanas indicadas. Harmon recibe periódicamente devoluciones de sus diversos proveedores. Se trata de instrumentos defectuosos por alguna razón o que se dañaron al ser enviados. Una vez hechas las reparaciones, las trompetas se devuelven al depósito de materiales listos para envío. Sobre la base de las devoluciones actuales y anticipadas, la compañía espera el siguiente programa de entradas al inventario:

Semana	8	9	10	11
Entradas programadas	12		6	9

Además de las entradas programadas, la compañía espera contar con 23 trompetas en inventario al final de la semana 7. El MPS de las trompetas se obtiene igualando el inventario disponible al final de la semana 7 con las entradas programadas para obtener la demanda neta predicha:

Semana	8	9	10	11	12	13	14	15	16	17
Demanda neta predicha	42	42	32	12	26	112	45	14	76	38

Habiendo determinado el MPS del producto final, debemos convertirlo en un programa de producción para los componentes en el siguiente nivel de la estructura del producto. Éstos son el ensamble de campana y el ensamble de la caja de válvulas. Consideremos primero el ensamble de campana. El primer paso consiste en traducir el MPS para las trompetas en un conjunto de requerimientos brutos por semana para el ensamble de campana. Como se emplea exactamente un ensamble de campana para cada trompeta, se trata del mismo que el MPS. El siguiente paso consiste en sustraer cualquier inventario disponible o entradas programadas para obtener los requerimientos netos (en este caso no hay ninguno). Después, estos requerimientos netos se trasladan hacia atrás en el tiempo por una cantidad igual al tiempo de demora, que es de dos semanas en el caso del ensamble de campana, para obtener los requerimientos trasladados. Finalmente, el algoritmo de ajuste de tamaño de lote se aplica a los reque-

[1] El lector inteligente sabrá que las válvulas y las correderas no son idénticas. De ahí que cada válvula y cada corredera deberían considerarse como artículos separados. Sin embargo, si acordamos que las válvulas y las correderas corresponden a grupos coincidentes de tres, nuestro enfoque es válido. Esto nos permite demostrar el efecto multiplicador cuando se necesitan varios componentes para un solo artículo final.

rimientos llevados a cabo por fases, para obtener la liberación de pedido planeada. Si suponemos una regla de producción de lote por lote, obtenemos los siguientes cálculos del MRP para el ensamble de campana:

Semana	6	7	8	9	10	11	12	13	14	15	16	17
Requerimientos brutos			42	42	32	12	26	112	45	14	76	38
Requerimientos netos			42	42	32	12	26	112	45	14	76	38
Requerimientos netos trasladados en el tiempo	42	42	32	12	26	112	45	14	76	38		
Requisiciones planeadas (lote por lote)	42	42	32	12	26	112	45	14	76	38		

Lote por lote implica que la cantidad de producción semanal es sencillamente el requerimiento neto trasladado en el tiempo. Una regla de producción lote por lote implica que no se lleva a cabo ningún inventario de un periodo a otro. Como más tarde veremos, la regla lote por lote rara vez es una regla de producción óptima. Las reglas de programación de la producción óptima y heurística se analizarán en la sección 7.2.

Los cálculos son esencialmente los mismos en el caso del ensamble de la caja de válvulas, salvo que el tiempo de entrega de la producción es de cuatro semanas en lugar de dos. Los cálculos para el ensamble de la caja de válvulas son:

Semana	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Requerimientos brutos					42	42	32	12	26	112	45	14	76	38
Requerimientos netos					42	42	32	12	26	112	45	14	76	38
Requerimientos netos trasladados en el tiempo	42	42	32	12	26	112	45	14	76	38				
Requisiciones planeadas (lote por lote)	42	42	32	12	26	112	45	14	76	38				

Ahora consideremos los cálculos del MRP para las válvulas. Supongamos que la compañía espera un inventario disponible de 186 válvulas al final de la semana 3 y una entrada de un proveedor de 96 válvulas al inicio de la semana 5. Se requieren tres válvulas para cada trompeta. (Nótese que las válvulas no son idénticas; de ahí que no sean intercambiables. Podríamos presentar tres conjuntos separados de cálculos MRP, pero resulta innecesario en virtud de que cada trompeta posee exactamente una válvula de cada tipo.) Obtenemos los requerimientos brutos para las válvulas multiplicando el programa de producción para el ensamble de la caja de válvulas por 3. Los requerimientos netos se calculan restando el inventario disponible y las entradas programadas. Los cálculos MRP para estas válvulas son los siguientes:

Semana	2	3	4	5	6	7	8	9	10	11	12	13
Requerimientos brutos			126	126	96	36	78	336	135	42	228	114
Entradas programadas				96								
Inventario disponible		186	60	30								
Requerimientos netos			0	0	66	36	78	336	135	42	228	114
Requerimientos netos trasladados en el tiempo		66	36	78	336	135	42	228	114			
Requisiciones planeadas (lote por lote)		66	36	78	336	135	42	228	114			

Los requerimientos netos se obtienen restando el inventario disponible y las entradas programadas de los requerimientos brutos. Como el inventario disponible de 186 en el periodo 3 excede el requerimiento total en el periodo 4, los requerimientos netos para el periodo 4 son 0. Las restantes 60 unidades ($186 - 126$) se trasladan al periodo 5. En el periodo 5 la entrada programada de 96 se añade al inventario inicial de 60 para obtener 156 unidades. Los requerimientos brutos para el periodo 5 son de 126; así que los requerimientos netos para el periodo 5 son 0, y las 30 unidades adicionales se trasladan al periodo 6. De ahí que los requerimientos netos resultantes para el periodo 6 sean $96 - 30 = 66$.

Los requerimientos netos se trasladan tres periodos hacia atrás con el fin de obtener los requerimientos netos trasladados y el programa de producción. Obsérvese que las válvulas se producen internamente. La entrada programada de 96 corresponde a los artículos defectuosos que se enviaron a reparación. Se requieren cálculos similares en el caso de los ensambles corredizos.

El ejemplo 7.1 representa los elementos esenciales del cálculo de la explosión. Nótese que, para el ejemplo, hemos supuesto que la regla de programación era lote por lote. Es decir, que en cada periodo la cantidad de producción es igual a los requerimientos netos para dicho periodo. Sin embargo, dicha política puede no ser óptima e incluso no factible. El programa requiere la entrega de 336 válvulas en la semana 9; supongamos que la planta puede producir solamente 200 válvulas en una semana: si ese fuera el caso, una regla de programación lote por lote no resultaría factible.

Problemas para la sección 7.1

1. Los modelos de control de inventarios analizados en los capítulos 4 y 5 a menudo se identifican como modelos de demanda *independiente* y el MRP con frecuencia se identifica como sistema de demanda *dependiente*. ¿Qué implican los términos *independiente* y *dependiente* en este contexto?
2. ¿Qué información se encuentra contenida en un diagrama de estructura del producto?
3. En el ejemplo de Harmon Music de esta sección, determine la requisición planeada en el caso de los ensambles corredizos. Supóngase una programación lote por lote.
4. La compañía Noname Computer construye una computadora designada modelo ICU2. Importa la tarjeta madre de la computadora de Taiwán, pero inserta los receptáculos para los chips y las tarjetas en su planta de Lubbock, Texas. Cada computadora requiere un total de 90 chips de 64K de memoria dinámica de acceso aleatorio (DRAM). Noname vende las computadoras con tres tarjetas agregadas y dos unidades de disco. La compañía compra tanto los chips DRAM como las unidades de disco a un proveedor externo. El diagrama de estructura del producto para la computadora ICU2 se muestra en la figura 7-6.

   Suponga que las demandas pronosticadas para la computadora en las semanas 6 a 11 son 220, 165, 180, 120, 75, 300. El inventario inicial de computadoras ensambladas la semana 6 será de 75, y el gerente de producción anticipa devoluciones de 30 en la semana 8 y 10 en la semana 10.

   a) Determine el MPS para las computadoras.

   b) Determine la requisición planeada para los tableros base suponiendo una regla de programación lote por lote.

   c) Determine el programa de pedidos externos de las unidades de disco.

5. Para el problema 4, suponga que Noname tiene 23 000 chips DRAM en existencia. Noname prevé recibir un lote de 3 000 chips la semana 3 de otra compañía que ha salido del

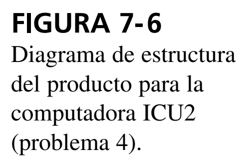
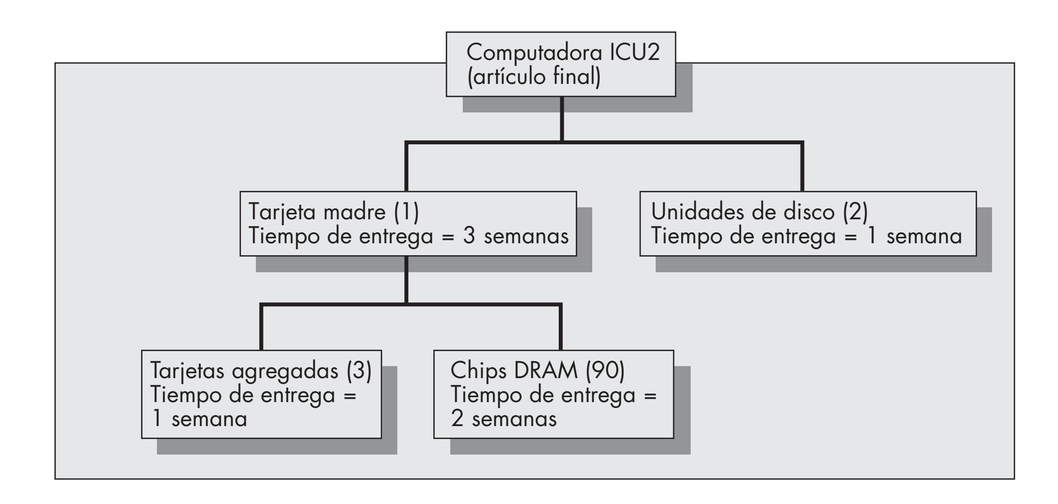


FIGURA 7-6
Diagrama de estructura del producto para la computadora ICU2 (problema 4).



mercado. En la actualidad, Noname compra los chips a dos vendedores, A y B. A vende los chips más baratos, pero no surtirá un pedido que rebase los 10 000 chips por semana.

a) Si Noname ha establecido una política de inventariar la menor cantidad de chips posible, ¿qué pedido debe hacer a los vendedores A y B las siguientes seis semanas?

b) Noname se ha dado cuenta de que no todos los chips DRAM comprados funcionan adecuadamente. Sobre la base de su experiencia en el pasado, calcula 8 por ciento de fallas en los chips comprados al vendedor A y 4 por ciento de fallas en los chips comprados al vendedor B. ¿Qué modificación en el programa de pedidos recomendaría usted con el fin de contrarrestar este problema?

6. Considérese el diagrama de estructura del producto de la figura 7-3. Suponga que el MPS para el artículo final en las semanas 10 a 17 es el siguiente:

Semana	10	11	12	13	14	15	16	17
Requerimientos netos	100	100	40	40	100	200	200	200

Suponga que la programación lote por lote se emplea en todo el proceso. También suponga que no existe inventario de entrada en el periodo 10 y no hay entradas programadas.

a) Determine la requisición planeada para el componente *A*.

b) Determine la requisición planeada para el componente *B*.

c) Determine la requisición planeada para el componente *C*. (Sugerencia: Observe que A y B requieren de C.)

7. ¿Qué alternativas existen para la programación lote por lote en cada nivel? Describa las posibles ventajas y desventajas de las técnicas de determinación de tamaños de lote.
8. Una de las entradas al sistema MRP es el pronóstico de demanda para el artículo final en el horizonte de planeación. Desde el punto de vista de la producción, ¿qué ventajas existen en el caso de un sistema de pronóstico que facilita la demanda (es decir que genera pronósticos relativamente constantes) en comparación con uno que permite mayor exactitud, pero que arroja pronósticos "con picos", que cambian significativamente de un periodo al siguiente?
9. Un artículo final posee el diagrama de estructura del producto de la figura 7-7.

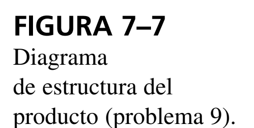
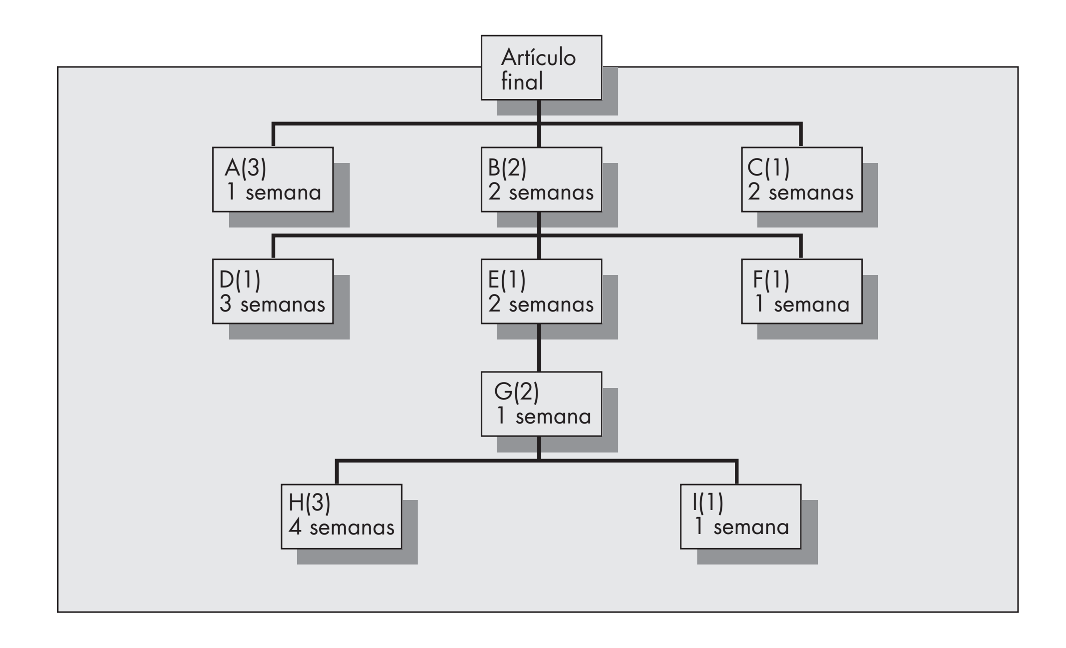


FIGURA 7–7
Diagrama
de estructura del
producto (problema 9).



a) Escriba el diagrama de estructura del producto en forma de lista indexada de materiales.
b) Suponga que el MPS para al artículo final es el siguiente:

Semana	30	31	32	33	34	35
MPS	165	180	300	220	200	240

Si la producción se programa lote a lote, determine la requisición planeada para el componente F.

c) A partir de los datos del inciso *b*), determine la requisición para el componente I.
d) A partir de los datos del inciso *b*), determine la requisición planeada para el componente H.

7.2 ESQUEMAS ALTERNATIVOS DE DETERMINACIÓN DE TAMAÑOS DE LOTE

En el ejemplo 7.1 supusimos que la regla de programación de la producción era lote por lote. Es decir que el número de unidades programadas para la producción cada periodo era el mismo que los requerimientos netos para dicho periodo. De hecho, se supuso esta política por conveniencia y comodidad. En general, no resulta óptima. El problema de hallar el mejor (o casi el mejor) programa de producción puede caracterizarse de la siguiente manera: tenemos un conjunto de demandas y costos de almacenamiento y preparación que varían con el tiempo. ¿Qué cantidades de producción reducirán al mínimo los costos totales de preparación y almacenamiento en el horizonte de planeación? Observe que ninguno de los métodos del capítulo 4 (que supone conocidas, aunque constantes, a las demandas) ni los del capítulo 5 (que supone demandas aleatorias) son adecuados.

En esta sección analizaremos varios métodos heurísticos (es decir, aproximados), que se utilizan para determinar tamaños de lote que son fácilmente incorporables en el cálculo MRP.

Determinación de tamaños de lote CEP

Para aplicar la fórmula CEP, necesitamos tres entradas: la tasa de demanda promedio, λ; la tasa de costo de almacenamiento, h; y el costo de preparación, K. Considérese el ensamble de caja de válvulas del ejemplo 7.1. Supongamos que la operación de preparación para la maquinaria empleada en esta operación de ensamblaje requiere el trabajo de dos empleados durante aproximadamente tres horas. Los trabajadores ganan un promedio de \$22 la hora. Esto se traduce en un costo de instalación de $(22)(2)(3) = \$132$.

La compañía utiliza un costo de almacenamiento basado en una tasa de interés anual de 22 por ciento. Cada ensamble de caja de válvulas cuesta a la compañía \$141.82 de materiales y el valor agregado por mano de obra. Así, los costos de almacenamiento ascienden a $(141.82)(0.22)/52 = \$0.60$ por ensamble de caja de válvulas a la semana.

La requisición planeada de una política de lote por lote requiere una producción semanal. Considere los costos de almacenamiento e instalación que se generan de las semanas 6 a 15 en las que se aplica esta política. Si adoptamos la convención de que los costos de almacenamiento se cargan al inventario cada semana, entonces el costo total de almacenamiento a lo largo del horizonte de la semana 10 equivale a cero. Ya que se lleva a cabo una preparación cada semana, el costo total de preparación generado durante el horizonte de planeación es $(132)(10) =$ \$1 320.

Este costo puede reducirse significativamente produciendo cantidades grandes con menor frecuencia. Como "primer recurso" podemos emplear la fórmula CEP para determinar una política de producción alternativa. La totalidad de los requerimientos netos trasladados durante las semanas 8 a 17 es de 439, para un promedio de 43.9 por semana. Si consideramos que $\lambda = 43.9$, $h = 0.60$ y $K = 132$, la fórmula CEP da como resultado

$$Q = \sqrt{\frac{2K\lambda}{h}} = \sqrt{\frac{(2)(132)(43.9)}{0.6}} = 139.$$

Si programamos la producción en tamaños de lote de 139 mientras se garantiza que todos los requerimientos netos se satisfagan, los cálculos MRP resultantes para el ensamble de la caja de válvulas son los siguientes:

Semana	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Requerimientos netos					42	42	32	12	26	112	45	14	76	38
Requerimientos trasladados en el tiempo	42	42	32	12	26	112	45	14	76	38				
Requisición planeada (CEP)	139	0	0	0	139	0	139	0	0	139				
Entregas planeadas					139	0	0	0	139	0	139	0	0	139
Inventario final					97	55	23	11	124	12	106	92	16	117

El inventario final de cada periodo se determina a partir de la fórmula

$$\text{Inventario final} = \text{Inventario inicial} + \text{Entregas planeadas} - \text{Requerimientos netos.}$$

Considere el costo de emplear dimensionamiento de lote con CEP en lugar de una determinación lote por lote. Durante los periodos 8 a 17 hay un total de cuatro instalaciones, lo cual da como resultado un costo total de preparación de (132)(4) = \$528. La forma más directa de calcular el costo de preparación consiste en simplemente acumular los inventarios finales para los 10 periodos y multiplicar por h. El inventario final acumulado es de $97 + 55 + 23 + \cdots + 117 = 653$. De esta manera, el costo total de preparación generado a lo largo de los 10 periodos es de $(0.63)(653) = \$391.80$. El costo total de preparación y almacenamiento cuando los tamaños de lote se calculan con la fórmula CEP es de $\$528 + \$391.80 = \$919.80$. Ésta constituye una mejora considerable en el costo de \$1,320 cuando se emplea una programación de la producción lote por lote. (No obstante, este ahorro no toma en cuenta el impacto del costo que el tamaño de lote puede tener sobre los niveles inferiores en la red de productos. Es posible, aunque poco probable, que en un sentido general, la política de lote por lote resulte más efectiva, desde el punto de vista de los costos, que la de CEP. Este punto se analizará con mayor profundidad en la sección 7.5.) Obsérvese que la aplicación de la CEP para establecer las cantidades de producción da como resultado un patrón completamente nuevo de requerimientos brutos para los ensambles corredizos y de válvulas un nivel abajo. En particular, los requerimientos brutos para las válvulas ahora son:

Semana	4	5	6	7	8	9	10	11	12	13
Requerimientos brutos	417	0	0	0	417	0	417	0	0	417

En el resto de la sección, analizaremos tres esquemas populares para la determinación de tamaños de lote cuando se conoce la demanda y el tiempo varía. Debe señalarse que el problema de determinar tamaños de lote sujetos a una demanda que varía con el tiempo aparece en contextos ajenos al MRP. Ilustraremos la forma en que estos métodos pueden vincularse al cálculo de la explosión MRP.

Heurística de Silver-Meal

La heurística de Silver-Meal (en honor de Halan Meal y Edward Silver) es un método de vanguardia que requiere la determinación del costo promedio por periodo como función del número de periodos que el pedido actual generará, y detener el cálculo cuando esta función se incremente.

Definamos $C(T)$ como el costo de almacenamiento e instalación promedio por periodo si el pedido actual genera los siguientes T periodos. Como antes, sean $(r_1, \ldots, r_n)$ los requerimientos en el horizonte del periodo n. Consideremos el periodo 1. Si producimos suficiente en el periodo 1 para satisfacer la demanda en ese mismo periodo, entonces incurrimos en el costo del pedido K. De ahí que

$$C(1) = K.$$

Si en el periodo 1 ordenamos lo suficiente para satisfacer la demanda en los periodos 1 y 2, entonces debemos almacenar r_2 para un periodo. Por lo tanto,

$$C(2) = (K + hr_2)/2.$$

Asimismo,

$$C(3) = (K + hr_2 + 2hr_3)/3$$

y, en general,

$$C(j) = (K + hr_2 + 2hr_3 + \cdots + (j - 1)hr_j)/j.$$

Una vez que $C(j) > C(j - 1)$, nos detenemos y establecemos $y_1 = r_1 + r_2 + \cdots + r_{j-1}$, e iniciamos el proceso de nuevo comenzando en el periodo j.

Ejemplo *7.2*

Un taller de maquinado emplea una heurística de Silver-Meal para programar la producción de lotes para cubiertas de computadora. En las siguientes cinco semanas, las demandas de cubiertas son $\mathbf{r} = (18, 30, 42, 5, 20)$. El costo de almacenamiento es de \$2 por cubierta a la semana, y el costo de preparación es de \$80. Determine los tamaños de lote recomendados.

Solución

Comenzando en el periodo 1:

$$C(1) = 80,$$
$$C(2) = [80 + (2)(30)]/2 = 70,$$
$$C(3) = [80 + (2)(30) + (2)(2)(42)]/3 = 102.67. \text{ Se detiene porque } C(3) > C(2).$$

Establezca $y_1 = r_1 + r_2 = 18 + 30 = 48$.

Comenzando en el periodo 3:

$$C(1) = 80,$$
$$C(2) = [80 + (2)(5)]/2 = 45,$$
$$C(3) = [80 + (2)(5) + (2)(2)(20)]/3 = 56.67. \text{ Se detiene.}$$

Establezca $y_3 = r_3 + r_4 = 42 + 5 = 47$.

Como el periodo 5 es el periodo final en el horizonte, no necesitamos comenzar el proceso de nuevo. Establecemos que $y_5 = r_5 = 20$. De ahí que la heurística de Silver-Meal dé como resultado la política $\mathbf{y} = (48, 0, 47, 0, 20)$. (Sugerencia: es posible llevar a cabo cálculos más eficientes si se observa que $C(j + 1) = [j/(j + 1)][C(j) + hr_{j+1}]$.)

Para demostrar que la heurística de Silver-Meal no siempre dará como resultado una solución óptima, consideremos el siguiente contraejemplo.

Ejemplo *7.3*

Sean $\mathbf{r} = (10, 40, 30)$, $K = 50$, y $h = 1$. La heurística de Silver-Meal da la solución $y = (50, 0, 30)$, pero la solución óptima es (10, 70, 0).

Para cerrar esta sección, observemos que Silver y Peterson (1985, p. 238) recomendaron condiciones bajo las cuales la heurística de Silver-Meal debería emplearse en lugar de la CEP. La condición se basa en la varianza de la demanda periódica: entre más alta sea la varianza, mayores mejoras proporcionará la heurística. Sin embargo, creemos que dada la tecnología actual para calcular y la facilidad con la que puede determinarse la solución heurística, los costos adicionales de los cálculos del uso de Silver-Meal (o uno de los dos siguientes métodos descritos) en lugar de la CEP son mínimos y no tienen gran importancia.

Costo unitario mínimo

La heurística de costo unitario mínimo (LUC, *least unit cost*) es similar al método de Silver-Meal, excepto que en lugar de dividir el costo en j periodos entre el número de periodos, j, lo dividimos entre la cantidad total de unidades demandadas a lo largo del periodo j, j, $r_1 + r_2 + \cdots + r_j$. Elegimos el horizonte de pedido que minimiza el costo por unidad de demanda, en lugar del costo por periodo.

Definimos $C(T)$ como el costo promedio de almacenamiento y preparación por unidad para un horizonte de pedido con T periodos. Así,

$$C(1) = K/r_1,$$
$$C(2) = (K + hr_2)/(r_1 + r_2),$$
$$\vdots$$
$$C(j) = [K + hr_2 + 2hr_3 + \cdots + (j-1)hr_j]/(r_1 + r_2 + \cdots + r_j).$$

Como en el caso de la heurística de Silver-Meal, este cálculo se detiene cuando $C(j) > C(j-1)$, y el nivel de producción se iguala a $r_1 + r_2 + \cdots + r_{j-1}$. Entonces se repite el proceso, comenzando en el periodo j y continuando hasta que se alcance el fin del horizonte de planeación.

Ejemplo *7.4* Suponga el mismo programa de requerimientos y costos dado en el ejemplo 7.2.

Comenzando en el periodo 1:

$$C(1) = 80/18 = 4.44,$$
$$C(2) = [80 + (2)(30)]/(18 + 30) = 2.92,$$
$$C(3) = [80 + (2)(30) + (2)(2)(42)]/(18 + 30 + 42) = 3.42.$$

Como $C(3) > C(2)$, nos detenemos y establecemos $y_1 = r_1 + r_2 = 48$.

Comenzando en el periodo 3:

$$C(1) = 80/42 = 1.90,$$
$$C(2) = [80 + (2)(5)]/(42 + 5) = 1.92.$$

Como $C(2) > C(1)$, nos detenemos y establecemos $y_3 = r_3 = 42$.

Comenzando en el periodo 4:

$$C(1) = 80/5 = 16,$$
$$C(2) = [80 + (2)(20)]/(5 + 20) = 4.8.$$

Como hemos alcanzado el final del horizonte, establecemos $y_4 = r_4 + r_5 = 5 + 20 = 25$. La solución obtenida por la heurística LUC es $\mathbf{y} = (48, 0, 42, 25, 0)$. Resulta interesante observar que la política obtenida por este método es diferente de la que se obtuvo para la heurística Silver-Meal. Resulta que el método de Silver-Meal proporciona la política óptima, con un costo de \$130, mientras que el LUC proporciona una política no óptima, con un costo de \$340.

Balanceo de periodos parciales

Otro método de aproximación para resolver este problema es el balanceo de periodos parciales. Aunque la técnica de Silver-Meal aparenta dar mejores resultados en un mayor número de casos, el balanceo de periodos parciales parece ser más frecuente en la práctica.

Este último método consiste en igualar el horizonte de pedido al número de periodos que más se ajuste al costo total de almacenamiento con el costo de preparación en dicho periodo. El horizonte de pedido que iguale exactamente los costos de preparación y almacenamiento difícilmente será un número entero de periodos (de ahí el nombre del método).

Ejemplo *7.5* De nuevo, considere el ejemplo 7.2. Comenzando en el periodo 1, encontramos que:

Horizonte de pedido	Costo total de almacenamiento
1	0
2	60
3	228

Como 228 excede el costo de instalación de 80, nos detenemos. Puesto que 80 está más cerca de 60 que de 228, el primer horizonte de pedido es de dos periodos. Es decir que $y_1 = r_1 + r_2 = 18 + 30 = 48$.

Iniciamos el proceso de nuevo en el periodo 3.

Horizonte del pedido	Costo total de almacenamiento
1	0
2	10
3	90

Hemos excedido los costos de instalación de 80, así que nos detenemos. Como 90 se encuentra más cerca de 80 que de 10, el horizonte del pedido es de tres periodos. Por lo tanto, $y_3 = r_3 + r_4 + r_5 = 67$. La solución del balanceo de periodos parciales completa es $y = (48, 0, 67, 0, 0)$, que es diferente a las soluciones de Silver-Meal y LUC. Esta solución es óptima, puesto que también posee un costo total de $310.

Los tres métodos analizados en esta sección son métodos heurísticos. Es decir que son métodos razonables basados en la estructura del problema, pero no necesariamente proporcionan la solución óptima. En el apéndice 7-A, se describe el algoritmo de Wagner-Whitin, el cual asegura una solución óptima al problema de planeación de la producción con demandas que varían con el tiempo. Aunque resulta tedioso resolverlo a mano, el algoritmo de Wagner-Whitin puede llevarse a cabo con facilidad en una computadora y es posible resolverlo con rapidez y eficiencia.

Problemas para la sección 7.2

10. Realice los cálculos MRP para las válvulas en el ejemplo de esta sección utilizando el programa de requerimientos brutos que se obtiene de la determinación del tamaño de lote CEP para los ensambles de cajas de válvulas. Suponga que $K = \$150$ y $h = 0.4$.
11. *a*) Determine la requisición planeada para las tarjetas madre del problema 4 suponiendo que se emplea la fórmula CEP para programar la producción. Considere que $K = \$180$ y $h = 0.40$.
    b) Con base en los resultados del inciso *a*, determine el programa de requerimientos brutos para los chips DRAM, que se piden a un proveedor externo. El costo del pedido es de $25.00 y el costo de almacenamiento es de $0.01 por chip cada semana. ¿Qué programa de pedidos al vendedor resulta si se emplea la fórmula CEP para determinar el tamaño del lote?
    c) Repita los cálculos del inciso *b* para los tableros añadidos. Utilice el mismo valor del costo de instalación y un costo de almacenamiento de 28 centavos por tablero a la semana.
12. *a*) Comente por qué la fórmula CEP puede arrojar malos resultados al determinar las requisiciones planeadas.
    b) Si la demanda predicha para el artículo final es la misma cada periodo, ¿dará como resultado la fórmula CEP un tamaño de lote óptimo en cada nivel de la estructura del producto? Explique.
13. El problema de determinar el tamaño de lote del ensamble de caja de válvulas descrito en el caso de la compañía Harmon Music, en la sección 7.2, se resolvió con la fórmula CEP. Determine el tamaño de lote para los 10 periodos con los siguientes métodos:
    a) Silver-Meal.
    b) Costo unitario mínimo.
    c) Balanceo de periodos parciales.
    d) ¿Cuál método de determinación de tamaños de lote da como resultado el costo más bajo para los 10 periodos?

14. Se ordena un solo artículo de inventario a un proveedor externo. La demanda anticipada de este artículo en los siguientes 12 meses es de 6, 12, 4, 8, 15, 25, 20, 5, 10, 20, 5, 12. Las existencias actuales de este artículo son de 4, y el inventario final debe ser de 8. Suponga un costo de almacenamiento de $1 por periodo y un costo de preparación de $40. Determine la política de pedidos para este artículo con base en:

    a) Silver-Meal.
    b) Costo unitario mínimo.
    c) Balanceo de periodos parciales.
    d) ¿Cuál método de determinación de tamaños de lote da como resultado el costo mínimo para los 12 periodos?

15. En el caso del contraejemplo (ejemplo 6.3), que demuestra que la heurística de Silver-Meal puede proporcionar una solución no óptima, ¿el costo unitario mínimo o el balanceo de periodos parciales proporcionan la solución óptima?

16. Comente las ventajas y desventajas de los siguientes métodos para determinar tamaños de lote en el contexto de un sistema de programación MRP: lote por lote, CEP, Silver-Meal, costo unitario mínimo y balanceo de periodos parciales.

17. Los requerimientos netos para montar la base de una lámpara de mesa las siguientes seis semanas son:

Semana	1	2	3	4	5	6
Requerimientos	335	200	140	440	300	200

    El costo de preparación para construir el ensamble de la base es de $200, y el costo de almacenamiento es de $0.30 por ensamble a la semana.

    a) ¿Qué tamaño de lote se obtiene de la fórmula CEP?
    b) Determine los tamaños de lote empleando la heurística Silver-Meal.
    c) Determine los tamaños de lote empleando la heurística del costo unitario mínimo.
    d) Determine los tamaños de lote empleando el balanceo de periodos parciales.
    e) Compare los costos de almacenamiento e instalación durante los seis periodos empleando las políticas que encontró en los incisos *a* al *d* con una política de lote por lote.

Los problemas del 18 al 22 se basan en el material que aparece en el apéndice 7-A.

18. Las demandas anticipadas para un horizonte de planeación de cuatro periodos son de 23, 86, 40 y 12. El costo de preparación es de $300 y el costo de almacenamiento es de $h = \$3$ por unidad por periodo.

    a) Enumere todas las políticas de requerimientos exactos, calcule los costos de preparación y almacenamiento para cada una y determine un programa de producción óptimo.
    b) Resuelva el problema mediante una programación dinámica regresiva.

19. Las demandas anticipadas para un horizonte de planeación de cinco periodos son de 14, 3, 0, 26 y 15. El inventario de inicio actual es de cuatro unidades, y el gerente de inventarios quisiera tener ocho unidades disponibles al final del horizonte de planeación. Suponga que $h = 1$ y $K = 30$. Determine el programa de producción óptimo. (Sugerencia: modifique las demandas del primero y último periodos para explicar las existencias iniciales y finales.)

20. Una pequeña empresa manufacturera que produce una línea de muebles de oficina requiere ruedas, con una tasa bastante constante de 75 por semana. El sistema MRP supone un horizonte de planeación de seis semanas. Suponga que los costos de preparación de la producción de las ruedas ascienden a $266 y los costos de almacenamiento son de $1 por rueda a la semana.

    a) Calcule la CEP y determine el número de periodos de demanda a los que ésta corresponde formando la razón (CEP)/(demanda por periodo). Sea T esta razón redondeada al

entero más próximo. Determine la política que produce ruedas una vez cada T periodos.

b) Utilizando un programa dinámico regresivo con $N = 6$ y $\mathbf{r} = (75, 75, \ldots, 75)$, determine la solución óptima. (Consulte el apéndice 7-A.) ¿Concuerda la respuesta con lo que se obtuvo en el inciso *a*?

21. *a*) Con base en los resultados del problema 20, sugiera una técnica aproximada para determinar tamaños de lote. ¿Bajo qué circunstancias se esperaría que este método diera buenos resultados?

b) Utilice este método para resolver el ejemplo 7A.2 (en el apéndice 7-A). ¿En qué porcentaje difiere la solución resultante de la óptima?

22. Resuelva el problema 17 utilizando el algoritmo de Wagner-Whitin (consulte el apéndice 7-A).

7.3 INCORPORACIÓN DE LOS ALGORITMOS PARA DETERMINAR TAMAÑOS DE LOTE EN EL CÁLCULO DE LA EXPLOSIÓN

Ejemplo *7.6*

Volvamos al ejemplo 7.1 relativo a la compañía Harmon Music y consideremos el impacto de la determinación del tamaño de lote al calcular la explosión. Primero consideremos el ensamble de caja de las válvulas. Los requerimientos netos trasladados para el ensamble de la caja de válvulas son:

Semana	4	5	6	7	8	9	10	11	12	13
Requerimiento neto trasladado en el tiempo	42	42	32	12	26	112	45	14	76	38

El costo de preparación del ensamble de caja de válvulas es de \$132, y el costo de almacenamiento de $h = \$0.60$ por ensamble cada semana. Determinaremos el tamaño de lote por la heurística de Silver-Meal.

Comenzando la semana 4:

$$C(1) = 132,$$

$$C(2) = \frac{132 + 0.6(42)}{2} = 78.6,$$

$$C(3) = \frac{132 + 0.6[42 + (2)(32)]}{3} = 65.2,$$

$$C(4) = \frac{132 + 0.6[42 + (2)(32) + (3)(12)]}{4} = 54.3,$$

$$C(5) = \frac{132 + 0.6[42 + (2)(32) + (3)(12) + (4)(26)]}{5} = 55.92.$$

Como $C(5) > C(4)$, concluimos los cálculos y establecemos que $y_4 = 42 + 42 + 32 + 12 = 128$.

Comenzando en la semana 8:

$$C(1) = 132,$$

$$C(2) = \frac{132 + 0.6(112)}{2} = 99.6,$$

$$C(3) = \frac{132 + 0.6[112 + (2)(45)]}{3} = 84.4,$$

$$C(4) = \frac{132 + 0.6[112 + (2)(45) + (3)(14)]}{4} = 69.6,$$

$$C(5) = \frac{132 + 0.6[112 + (2)(45) + (3)(14) + (4)(76)]}{5} = 92.16.$$

Por lo tanto, $y_8 = 26 + 112 + 45 + 14 = 197$.

La producción ocurre enseguida, en la semana 12. Resulta fácil mostrar que $y_{12} = 76 + 38 = 114$.

A continuación se presenta un resumen de los cálculos MRP con la heurística de Silver-Meal (S-M) para determinar tamaños de lote para el ensamble de la caja de las válvulas:

Semana	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Requerimientos netos					42	42	32	12	26	112	45	14	76	38
Requerimientos netos trasladados en el tiempo	42	42	32	12	26	112	45	14	76	38				
Requisición planeada (S–M)	128	0	0	0	197	0	0	0	114	0				
Entregas planeadas					128	0	0	0	197	0	0	0	114	0
Inventario final					86	44	12	0	171	59	14	0	38	0

Resulta interesante comparar el costo de almacenamiento y de preparación de la política en la que se empleó la heurística de Silver-Meal con las soluciones anteriores en las que se utilizó una técnica de lote por lote y la fórmula CEP. Existen exactamente tres preparaciones en nuestra solución, lo cual da como resultado un costo total de preparación de (132)(3) = \$396. La suma de los inventarios finales cada semana es de $86 + 44 + \cdots + 38 + 0 = 424$. El costo total de la solución Silver-Meal para este ensamble asciende a \$650.40. Compare esto con los costos de \$1,320 lote por lote y con la solución CEP.

También es interesante observar cuál hubiera sido el resultado de haber aplicado el algoritmo de Wagner-Within para encontrar la solución óptima. La solución óptima para este problema resulta ser $y_4 = 154$, $y_9 = 171$ y $y_{12} = 114$, con un costo total de \$610.20, que es sólo una mejora ligera sobre la heurística de Silver-Meal.

Ahora consideraremos cómo la emisión de requisiciones planeadas para el ensamble de la caja de válvulas influye en la programación de los componentes de nivel inferior. En particular, si tomamos en cuenta los cálculos MRP para las válvulas y suponemos que el tamaño de lote para las mismas se determina por medio de la heurística de Silver-Meal, obtenemos lo siguiente:

Semana	1	2	3	4	5	6	7	8	9	10	11	12	13
Requerimientos brutos				384	0	0	0	591	0	0	0	342	0
Entradas programadas					96								
Inventario disponible			186	0	96	96	96	0					
Requerimientos netos				198	0	0	0	495	0	0	0	342	0
Requerimientos netos trasladados en el tiempo	198	0	0	0	495	0	0	0	342	0			
Requisiciones planeadas (S–M)	198	0	0	0	495	0	0	0	342	0			

Observe que en este resumen de cálculos las entradas programadas en el periodo 5 deben mantenerse hasta el periodo 8 antes de que se puedan emplear para compensar la demanda. Esto da como resultado cero requerimientos brutos en los periodos del 5 al 8.

El cálculo de la emisión de requisiciones planeadas se basa en un costo de preparación de \$80 y un costo de almacenamiento de \$0.07 por válvula a la semana. Resulta significativo que, en este caso, la heurística Silver-Meal arrojó una regla de producción lote por lote. Esto resulta del patrón de demanda irregular causado por el tamaño de lote aplicado a un nivel superior de la estructura del producto. Tanto Silver-Meal como Wagner-Whitin proporcionan los mismos resultados para este ejemplo.

Problemas para la sección 7.3

23. Si hubiera que resolver el ejemplo 7.3 utilizando el algoritmo de Wagner-Whitin (expuesto en el apéndice 7-A), obtendríamos (154, 0, 0, 0, 0, 171, 0, 0, 114, 0) como requisición planeada para el ensamble de la caja de válvulas. ¿Cuáles son las requisiciones planeadas para las válvulas?
24. Considere el ejemplo de la compañía Noname Computer planteado en el problema 4. Suponga que el costo de preparación de la producción de las tarjetas madre es de \$180 y que el costo de almacenamiento es $h = \$0.40$ por tarjeta madre cada semana. Aplicando un balanceo de periodos parciales, determine las requisiciones planeadas para las tarjetas madre

y el programa de requerimientos brutos resultante para los chips DRAM. (Sugerencia: Utilice la demanda neta para las computadoras después de tomar en cuenta el inventario inicial y las devoluciones.)

25. En el caso del problema 6, suponga que el costo de instalación para ambos componentes, A y B, es de \$100, y que los costos de almacenamiento son, respectivamente, de \$0.15 y \$0.25 por componente cada semana. Con la ayuda del algoritmo de Silver-Meal, determine las requisiciones planeadas para ambos componentes, A y B, y los programas de requerimientos brutos resultantes para los componentes C, D y E.

7.4 DETERMINACIÓN DE TAMAÑOS DE LOTE CON RESTRICCIONES DE CAPACIDAD

Consideremos una variable del problema analizado en la sección 7.3. Supongamos que, además de los requerimientos conocidos ($r_1, \ldots, r_n$) en cada periodo, también existen capacidades de producción ($c_1, \ldots, c_n$). Por consiguiente, queremos determinar las cantidades de producción óptima ($y_1, \ldots, y_n$) sujetas a las restricciones $y_i \leq c_i$, para $1 \leq i \leq n$.

La introducción de restricciones de capacidad vuelve más realista el problema. Como los algoritmos para determinar tamaños de lote pueden incorporarse al sistema de planeación MRP, las capacidades de producción resultarán una parte importante de cualquier solución realizable. Sin embargo, éstas también hacen el problema más complejo. El resultado supuestamente obvio de que las políticas óptimas siempre ordenan requerimientos exactos ya no es válido. La determinación de políticas óptimas es difícil y consume demasiado tiempo, además de que probablemente no resulte práctica para la mayoría de los problemas reales.

Incluso obtener una solución factible puede no ser tan evidente. Consideremos nuestro ejemplo simple de cuatro periodos con un vector $\mathbf{r} = (52, 87, 23, 56)$, pero ahora supongamos que la capacidad de producción en cada periodo es $\mathbf{c} = (60, 60, 60, 60)$. Primero debemos determinar si el problema es factible; es decir, si existe por lo menos una solución. A simple vista, parece que el problema puede resolverse, puesto que el requerimiento total a lo largo de los cuatro periodos es de 218 y la capacidad total es de 240. Pero este problema no es factible; lo más que puede producirse en los primeros dos periodos es 120, pero los requerimientos para dichos periodos suman 139.

Tenemos la siguiente condición de factibilidad:

$$\sum_{i=1}^{j} c_i \geq \sum_{i=1}^{j} r_i \qquad \text{para } j = 1, \ldots, n.$$

Aun cuando la condición de viabilidad se satisface, no resulta evidente la forma de encontrar una solución factible. Considere el siguiente ejemplo:

Ejemplo ***7.7***

$$\mathbf{r} = (20, 40, 100, 35, 80, 75, 25),$$
$$\mathbf{c} = (60, 60, 60, 60, 60, 60, 60).$$

Al comprobar la viabilidad, tenemos que:

$r_1 = 20,$ $\quad c_1 = 60;$

$r_1 + r_2 = 60,$ $\quad c_1 + c_2 = 120;$

$r_1 + r_2 + r_3 = 160,$ $\quad c_1 + c_2 + c_3 = 180;$

$r_1 + r_2 + r_3 + r_4 = 195,$ $\quad c_1 + c_2 + c_3 + c_4 = 240;$

$r_1 + r_2 + r_3 + r_4 + r_5 = 275,$ $\quad c_1 + c_2 + c_3 + c_4 + c_5 = 300;$

$r_1 + r_2 + r_3 + r_4 + r_5 + r_6 = 350,$ $\quad c_1 + c_2 + c_3 + c_4 + c_5 + c_6 = 360;$

$r_1 + r_2 + r_3 + r_4 + r_5 + r_6 + r_7 = 375,$ $\quad c_1 + c_2 + c_3 + c_4 + c_5 + c_6 + c_7 = 420.$

La prueba de factibilidad se satisface, así sabemos que existe por lo menos una solución factible. Sin embargo, no resulta tan evidente la forma en que podemos hallar una. La programación lote por lote no funcionará como consecuencia de las restricciones de la capacidad en los periodos 3, 5 y 6.

Expondremos una técnica de aproximación para la determinación de tamaños de lote a fin de obtener una solución inicial factible. El método consiste en trasladar la demanda desde periodos en los que la demanda excede la capacidad hasta periodos anteriores en los que existe una capacidad excesiva. Este proceso se repite en cada periodo cuando la demanda excede la capacidad hasta que se construye un nuevo programa de requerimientos donde el procedimiento de lote por lote resulta factible. En el ejemplo, el primer periodo en el que la demanda excede la capacidad es el periodo 3. Sustituimos r_3 por c_3. La diferencia de 40 unidades ahora debe redistribuirse en los periodos 1 y 2. Consideremos el primer periodo anterior, que es el periodo 2. Hay 20 unidades excedentes en la capacidad en el periodo 2, que absorbemos. Aún tenemos 20 unidades de demanda del periodo 3 que no se han tomado en cuenta; éstas se añaden al requerimiento para el periodo 1. Al resumir los resultados hasta este punto, tenemos lo siguiente:

$$\mathbf{r}' = (\overset{40}{\cancel{20}}, \overset{60}{\cancel{40}}, \overset{60}{\cancel{100}}, 35, 80, 75, 25),$$
$$\mathbf{c} = (60, 60, 60, 60, 60, 60, 60).$$

El siguiente periodo en que la capacidad se rebasa es el periodo 5. El excedente en la demanda de 20 unidades puede trasladarse al periodo 4. Finalmente, las 15 unidades de excedente en la demanda en el periodo 6 pueden distribuirse en los periodos 4 (5 unidades) y 1 (10 unidades). La condición de factibilidad garantiza que este proceso conduce a una solución factible.

Esto conduce a:

$$\mathbf{r}' = (\overset{\overset{50}{\cancel{40}}}{\cancel{20}}, \overset{60}{\cancel{40}}, \overset{60}{\cancel{100}}, \overset{\overset{60}{\cancel{55}}}{\cancel{35}}, \overset{60}{\cancel{80}}, \overset{60}{\cancel{75}}, 25),$$
$$\mathbf{c} = (60, 60, 60, 60, 60, 60, 60).$$

Por lo tanto, el programa de requerimientos modificado que se obtiene es:

$$\mathbf{r}' = (50, 60, 60, 60, 60, 60, 25).$$

Si establecemos que $\mathbf{y} = \mathbf{r}'$, se obtiene una solución factible para el problema original.

Etapa de mejora

Tenemos que la programación lote por lote para el programa de requerimientos modificado **r´** resulta viable en el caso de problema original. Enseguida nos interesa ver si podemos descubrir alguna forma de mejora; es decir, otra política viable de menor costo. Existe una variedad razonable de reglas de mejora que podemos emplear. Utilizaremos la siguiente.

Para cada lote que se programe, comenzando a partir del último y regresando al inicial, se determina si resulta más barato producir las unidades que conforman dicho lote trasladando la producción a periodos anteriores de exceso de capacidad. Si se elimina un lote, se reduce el costo de instalación en dicho periodo a cero, pero el traslado de la producción a periodos anteriores incrementa el costo de almacenamiento. Este cambio se realiza sólo si el costo de almacenamiento adicional es menor que el costo de instalación. Ilustremos el proceso con un ejemplo.

Ejemplo *7.8* Suponga que $K = \$450$ y $h = \$2$.

$$\mathbf{r} = (100, 79, 230, 105, 3, 10, 99, 126, 40),$$
$$\mathbf{c} = (120, 200, 200, 400, 300, 50, 120, 50, 30).$$

Si se calcula la suma acumulada de requerimientos y capacidades para cada periodo, resulta fácil ver que el problema es factible. Sin embargo, como los requerimientos exceden las capacidades en algunos periodos, una programación lote por lote no resulta factible. Trasladamos los excedentes en la demanda a periodos anteriores con el fin de obtener el programa de requerimientos modificado $\mathbf{r}' = (100, 109, 200, 105, 28, 50, 120, 50, 30)$.

La programación lote por lote para el programa de requerimientos modificado $\mathbf{r}'$ es factible para el problema original. La solución factible inicial requiere nueve preparaciones a un costo total de preparación de $9 \times 450 = \$4{,}050$. El costo de almacenamiento de la política inicial es $2(0 + 30 + 0 + 0 + 25 + 65 + 86 + 10) = \$432$.

Con el fin de ejecutar la etapa de mejora es conveniente ordenar los datos en una tabla.

	1	2	3	4	5	6	7	8	9
r′	100	109	200	105	28	50	120	50	30
c	120	200	200	400	300	50	120	50	30
y	100	109	200	105	28	50	120	50	30
Exceso de capacidad	20	91	0	295	272	0	0	0	0

Comenzando desde el último periodo, consideremos el lote final de 30 unidades. Hay suficiente exceso de capacidad en periodos anteriores para pensar en trasladar este lote. El último periodo donde puede programarse este lote es el periodo 5. El costo extra de almacenamiento generado por el traslado es de $2 \times 30 \times 4 = \$240$. Ya que esto es más barato que el costo de instalación de \$450, llevamos a cabo el cambio, incrementamos y_5 de 28 a 58 y reducimos el exceso de capacidad en el periodo 5 de 272 a 242.

Ahora consideremos el lote de 50 unidades programado el periodo 8. Este lote también se puede trasladar al periodo 5 con un costo de almacenamiento adicional resultante de $2 \times 50 \times 3 = \$300$. De nuevo resulta más barato que el costo de preparación, así que podemos realizar el cambio. En este momento tenemos que $y_5 = 108$, y el excedente de capacidad en el periodo 5 se reduce a 192.

Los cálculos se resumen en la tabla de la siguiente manera:

	1	2	3	4	5	6	7	8	9
r′	100	109	200	105	28	50	120	50	30
c	120	200	200	400	300	50	120	50	30
					108				
					~~58~~			0	0
y	100	109	200	105	~~28~~	50	120	~~50~~	~~30~~
					192				
					~~242~~				
Exceso de capacidad	20	91	0	295	~~272~~	0	0	0	0

Ahora consideremos el lote de 120 unidades programado en el periodo 7. En este momento, aún contamos con 192 unidades de exceso de capacidad en el periodo 5. El costo de almacenamiento adicional por desplazar el lote de 120 del periodo 7 al periodo 5 es de $2 \times 120 \times 2 = \$480$. Esto rebasa el costo de preparación de \$450, así que no hacemos el cambio.

Evidentemente resulta ventajoso trasladar el lote de 50 unidades del periodo 6 al periodo 5, lo cual reduce el exceso de capacidad en el periodo 5 a 142 e incrementa el tamaño del lote el periodo 5 de 108 a 158. Si se hace esto se obtienen los siguientes resultados:

	1	2	3	4	5	6	7	8	9
r′	100	109	200	105	28	50	120	50	30
c	120	200	200	400	300	50	120	50	30
					158				
					~~108~~				
					~~58~~			0	0
y	100	109	200	105	~~28~~	0	120	~~50~~	~~30~~
					142				
					~~192~~				
					~~242~~				
Exceso de capacidad	20	91	0	295	~~272~~	0	0	0	0

En este momento parece que hemos concluido. Sin embargo, hay suficiente capacidad en el periodo 4 como para trasladar todo el lote de 158 unidades del periodo 5 al periodo 4. El costo adicional de preparación por hacer esto es de $2 \times 158 = \$316$. Como lo anterior resulta más barato que el costo de preparación, hacemos el cambio.

Si resumimos estos cálculos en la tabla, obtenemos:

	1	2	3	4	5	6	7	8	9
r′	100	109	200	105	28	50	120	50	30
c	120	200	200	400	300	50	120	50	30
					0				
					~~158~~				
					~~108~~				
				263	~~58~~	0		0	0
y	100	109	200	~~105~~	~~28~~	~~50~~	120	~~50~~	~~30~~
					300				
					~~142~~				
					~~192~~				
				137	~~242~~	50		50	30
Exceso de capacidad	20	91	0	~~295~~	~~272~~	~~0~~	0	~~0~~	~~0~~

En este momento no es posible ningún otro traslado de lote. La solución que hemos obtenido es la siguiente:

$$\mathbf{y} = (100, 109, 200, 263, 0, 0, 120, 0, 0)$$

para el programa de requerimientos original

$$\mathbf{r} = (100, 79, 230, 105, 3, 10, 99, 126, 40).$$

Calcularemos el costo de esta solución y lo compararemos con el de la solución factible original. Hay cinco instalaciones con un costo total de $5 \times 450 = \$2{,}250$. El costo de instalación es de $2(0 + 30 + 0 + 158 + 155 + 145 + 166 + 40 + 0) = 2 \times 694 = \$1{,}388$. El costo total de esta política es de \$3,638, comparado con \$4,482 en el caso de la política factible inicial. En este ejemplo, la etapa de mejora dio como resultado una reducción del costo cercana al 20 por ciento.

Problemas para la sección 7.4

26. Considere el ejemplo de la sección 7.2 relacionado con la programación de la producción del ensamble de la caja de válvulas.

    a) Suponga que la capacidad de producción cualquier semana es de 100 cajas de válvulas. Si se aplica el algoritmo explicado en esta sección, determine las requisiciones planeadas para las cajas de válvulas.

    b) ¿Qué programa de requerimientos brutos para las válvulas proporciona el tamaño de lote que se obtuvo en el inciso *a*?

    c) Suponga que la capacidad de producción en el caso de las válvulas es de 200 válvulas a la semana. ¿Resulta factible el programa de requerimientos brutos del inciso *b*? Si no, sugiera una modificación en las requisiciones planeadas que se calcularon en el inciso *a* que dé como resultado un programa de requerimientos brutos factible para las válvulas.

27. Resuelva el problema 14 suponiendo un pedido mínimo de 20 por mes.

28. *a*) Resuelva el problema 17 suponiendo las siguientes capacidades de producción:

Semana	1	2	3	4	5	6
Capacidad	600	600	600	400	200	200

    b) En términos porcentuales, ¿qué tan mayor es el costo total de preparación y almacenamiento en el caso capacitado si se compara con las soluciones obtenidas en los incisos *b*, *c* y *d* del problema 17?

29. El método de reprogramación de la producción de un lote en uno o más de los periodos anteriores si el incremento en el costo de almacenamiento es menor que el costo de preparación también se puede emplear cuando no existen restricciones de capacidad. Este método constituye una técnica heurística alternativa de programación por lotes para el problema no capacitado. En el caso del problema 14, se comienza con una política de lote por lote y se considera el traslado regresivo de lotes, como lo hicimos en esta sección, comenzando con el periodo final y terminando con el primer periodo. Compare los costos totales de la política resultante con los de las políticas deducidas en el problema 14.

7.5 DESVENTAJAS DEL MRP

El MRP es un sistema de producción cerrada con dos entradas principales: 1) el programa maestro de producción para el artículo final y 2) las relaciones entre los diversos componentes, módulos y subensambles que conforman el proceso de producción para el artículo final. El método es lógico y evidentemente sensible a la programación de los tamaños de lote de la producción. Sin embargo, muchos de los supuestos no son realistas. En esta sección analizaremos algunos de dichos supuestos, los problemas que resultan como consecuencia de éstos y los medios para superar estos problemas.

Incertidumbre

Bajo el MRP se encuentra el supuesto de que toda la información que se requiere se conoce con certeza. Sin embargo, existen incertidumbres. Las dos fuentes fundamentales de incertidumbre son los pronósticos de ventas futuras del artículo final y la aproximación de los tiempos de abastecimiento de un nivel a otro. La incertidumbre del pronóstico normalmente implica que la satisfacción de la demanda probablemente sea diferente del pronóstico para dicha demanda. En el contexto de la planeación de la producción, esto también podría implicar que los pronósticos actualizados de las demandas futuras difieren de los pronósticos anteriores para dichas demandas. Los pronósticos deben revisarse cuando se acepten nuevos pedidos, se cancelen pedidos anteriores, o cuando esté disponible nueva información sobre el mercado. Esto tiene dos consecuencias en el sistema MRP. Una es que *todas* las decisiones sobre los tamaños de lote tomadas en la última actividad del sistema podrían ser incorrectas, y, algo más problemático, las decisiones anteriores que se tomen durante el proceso de producción pueden estar equivocadas.

El análisis de los modelos de inventario estocásticos del capítulo 5 mostró que una política óptima incluía una existencia de seguridad como protección contra la incertidumbre de la demanda. Es decir, haríamos pedidos hasta un nivel que exceda la demanda esperada. La misma lógica puede aplicarse a los sistemas MRP. La forma en que se transmite la incertidumbre a través de un sistema de producción complejo con muchos niveles no se entiende del todo. Por ello, no se recomienda incluir existencias de seguridad en todos los niveles del sistema. En su lugar, empleando los métodos del capítulo 5, se pueden construir niveles adecuados de seguridad en los pronósticos del artículo final. Éstos se transmiten automáticamente a través del sistema a niveles inferiores por medio del cálculo de la explosión.

Ejemplo *7.9*

Considere el ejemplo 7.1 relacionado con la compañía Harmon Music. Suponga que la empresa pretende incorporar incertidumbre en los pronósticos de demanda para las semanas de la 8 a la 17. Con base en el historial de las ventas de trompetas que conserva la firma, un analista se da cuenta de que la razón de la desviación estándar del error de pronóstico con respecto a la demanda media cada semana se encuentra cerca de 0.3.[2] Además, la demanda semanal se aproxima por medio de la distribución normal. Harmon decide producir suficientes trompetas para satisfacer la demanda cada semana con una probabilidad de .90. (Según la terminología del capítulo 5, esto significa que Harmon emplea un nivel

[2] En símbolos, esto se representa como σ/μ, que recibe el nombre de coeficiente de variación.

de servicio Tipo 1 del 90 por ciento para las trompetas.) La existencia de seguridad tiene la forma σz, donde *z* el punto adecuado en la tabla normal. En este caso, $z = 1.28$. Si se incorpora la existencia de seguridad en los pronósticos de demanda, obtenemos:

Semana	8	9	10	11	12	13	14	15	16	17
Demanda predicha (μ)	77	42	38	21	26	112	45	14	76	38
Desviación estándar (σ)	23.1	12.6	11.4	6.3	7.8	33.6	13.5	4.2	22.8	11.4
Demanda media más existencia de seguridad ($\mu + \sigma z$)	107	58	53	29	36	155	62	19	105	53

Por supuesto, ésta no es la única forma de calcular la existencia de seguridad. Existen alternativas en las que se emplea un criterio de servicio Tipo 2 o un modelo de costos por faltantes en lugar de un modelo de niveles de servicio. El siguiente paso consiste en generar los ingresos programados y el inventario disponible anticipado para llegar a un MPS revisado para las trompetas. El cálculo de la explosión ahora se realiza como antes, excepto que la existencia de seguridad incluida en el MPS revisado se transmitirá automáticamente a los ensambles de nivel inferior.

Los tiempos de demora de seguridad se emplean para compensar la incertidumbre de los tiempos de suministro de la producción en sistemas MRP. Dicho de manera simple, esto significa que los cálculos del tiempo que se requiere para completar un grupo de producción en un nivel y transportarlo al siguiente se multiplicarían por algún factor de seguridad. Si se empleara un factor de seguridad de 1.5 en la compañía Harmon Music, los tiempos de entrega de los componentes se revisarían de la siguiente manera: ensamble de la campana, 3 semanas; ensamble de la caja de válvulas, 6 semanas; ensambles corredizos, 3 semanas; válvulas, 4.5 semanas. Desde un punto de vista conceptual, los tiempos de entrega de seguridad tienen sentido si la incertidumbre esencial se encuentra en los tiempos de producción de un nivel al siguiente, y las existencias de seguridad tienen sentido si la incertidumbre esencial se encuentra en el pronóstico de la demanda de cada artículo final. En la práctica, ambas fuentes de incertidumbre por lo general se hallan presentes y se emplea cierta combinación de ambas existencias de seguridad y tiempos de entrega de seguridad.

Planeación de capacidad

Otro tema importante, que MRP aún no analiza explícitamente, tiene que ver con la capacidad de las instalaciones de producción. El método de determinación de lote con capacidad que se analizó antes tendrá que ver con las capacidades de producción en un nivel del sistema, pero no resolverá la cuestión de la capacidad total. El problema consiste en que, incluso si los tamaños de lote en determinado nivel no exceden las capacidades de producción, no hay garantía de que cuando estos tamaños de lote se conviertan en requerimientos brutos en un nivel inferior, estos requerimientos también puedan satisfacerse con la capacidad existente. Es decir que un programa de producción factible en cierto nivel puede dar como resultado un programa de requerimientos no factible en un nivel inferior.

La planeación de los requerimientos de capacidad (CRP, *capacity requirements planning*) es el proceso por medio del cual los requerimientos de capacidad en un centro de trabajo se calculan por medio de la salida de las liberaciones de pedidos planeados MRP. Si las requisiciones planeadas dan como resultado un programa de requerimientos inviable, hay varias acciones correctivas posibles. Una consiste en el tiempo extra estipulado en los lugares donde existan cuellos de botella. Otra consiste en revisar el MPS de manera tal que las requisiciones planeadas en niveles inferiores se logren con la capacidad actual del sistema. Ésta constituye una forma engorrosa de resolver el problema, pues requiere un proceso repetitivo de ensayo y error entre el CRP y el MRP.

Como en el ejemplo del CRP, consideremos la manufactura de la trompeta del ejemplo 7.1, así como en el resto del capítulo. Supongamos que las válvulas se fabrican en tres centros de trabajo: 100, 200 y 300. En el centro de trabajo 100, el latón fundido se vacía y se le da la forma de la válvula. En el centro de trabajo 200 se perforan los agujeros en las posiciones apropiadas en las válvulas (hay tres configuraciones de agujero, dependiendo del número de la válvula: 1, 2 o 3). Por último, en el centro de trabajo 300 la válvula se pule y la superficie se nivela adecuadamente para garantizar que la válvula no se pegue en la operación.

La siguiente tabla contiene un resumen de la información adecuada para los centros de trabajo:

Centro de trabajo	Tiempo de trabajador requerido para producir una unidad (horas/unidad)	Producción en máquina (unidades/días)
100	0.1	120
200	0.25	100
300	0.15	160

De acuerdo con esta información, se requeriría un total de seis minutos (0.1 horas) de tiempo de trabajo para producir una sola válvula en el centro de trabajo 100, y el equipo disponible puede soportar una cantidad máxima de 120 válvulas por día. Consideremos las requisiciones planeadas que se obtuvieron, para el caso de las válvulas, con la regla de programación por lotes de Silver-Meal en la sección 7.3:

Semana	2	3	4	5	6	7	8	9	10	11
Requisiciones planeadas (S–M)	198	0	0	0	495	0	0	0	342	0

Esta emisión planeada de requisiciones se traduce en los siguientes requerimientos de capacidad en los tres centros de trabajo.

Semana	2	3	4	5	6	7	8	9	10	11
Requerimientos de tiempo de mano de obra (horas):										
Centro de trabajo 100	19.8	0	0	0	49.5	0	0	0	34.2	0
Centro de trabajo 200	49.5	0	0	0	123.75	0	0	0	85.5	0
Centro de trabajo 300	29.7	0	0	0	72.25	0	0	0	51.3	0
Requerimiento de tiempo de máquina (días):										
Centro de trabajo 100	1.65	0	0	0	4.125	0	0	0	2.85	0
Centro de trabajo 200	1.98	0	0	0	4.95	0	0	0	3.42	0
Centro de trabajo 300	1.24	0	0	0	3.09	0	0	0	2.14	0

Los requerimientos de capacidad muestran si las requisiciones planeadas obtenidas con el MRP son factibles. Por ejemplo, supongamos que el requerimiento de 123.75 horas de mano de obra la semana 6 en el centro de trabajo 200 rebasa la capacidad de este centro de trabajo. Esto significa que el tamaño de lote actual no es factible y se requiere alguna acción correctiva. Una posibilidad consiste en dividir el lote programado para la semana 6 produciendo alguna parte de este en una semana anterior. Otra posibilidad consiste en ajustar el tamaño del lote para el ensamble de caja de válvulas en el siguiente nivel superior de la estructura del producto con el fin de acomodar las restricciones de la capacidad en el nivel actual. En cualquier caso, se requieren cambios sustanciales en el plan de producción inicial.

Estos ejemplos sugieren una interesante especulación. ¿No tendría más sentido determinar dónde ocurren los cuellos de botella *antes* de intentar explotar el MRP a través de varios niveles del sistema? De esta manera, se podría encontrar un plan de producción factible que cumpla con las restricciones de capacidad. Entonces pueden tomarse en cuenta refinamientos adicionales.

Horizontes progresivos y nerviosismo del sistema

Hasta ahora, nuestra perspectiva del MRP consiste en que se trata de un sistema estático. Dados los requerimientos para los artículos finales en un horizonte de planeación específico, pueden determinarse tanto los tiempos como los tamaños de los lotes de producción para los componentes de nivel inferior. No obstante, en la práctica, el entorno de planeación de la producción es dinámico. Puede ser necesario volver a ejecutar el sistema MRP cada periodo y volver a evaluar las decisiones de producción. A menudo sucede que sólo es necesario tomar las decisiones referentes a tamaños de lote para el periodo de planeación actual. Empleamos el término *horizonte progresivo* para referirnos a la situación en la que sólo se implementa la decisión

en el primer periodo de decisión de un problema de N periodos. El problema completo, para N periodos, se vuelve a correr en cada periodo para determinar una nueva decisión de primer periodo.

Cuando se emplean los horizontes progresivos, el horizonte de planeación deber ser suficientemente grande para garantizar que la decisión del primer periodo no cambie. Desafortunadamente en ciertos esquemas de demanda relativos a la decisión del primer periodo no permanece constante aún en el caso de horizontes de planeación grandes. Consideremos el siguiente ejemplo (de Carlson, Beckman y Kropp, 1982).

Ejemplo *7.10*

Suponga que la demanda se rige por el esquema cíclico 190, 210, 190, 210, 190, ... En un horizonte de planeación de cinco periodos, el programa de requerimientos para los periodos 1 a 5 es el siguiente:

$$\mathbf{r} = (190, 210, 190, 210, 190).$$

Además, suponga que $h = 1$ y $K = 400$. La solución óptima para este problema obtenida a partir del algoritmo de Wagner-Whitin es

$$\mathbf{y} = (190, 400, 0, 400, 0).$$

Sin embargo, suponga que el horizonte de planeación se elige de seis periodos en lugar de cinco. El programa de requerimientos para un horizonte de planeación de seis periodos es

$$\mathbf{r} = (190, 210, 190, 210, 190, 210).$$

La solución óptima en este caso es

$$\mathbf{y} = (400, 0, 400, 0, 400, 0).$$

Es decir, la producción del primer periodo ha cambiado de 190 a 400. Si se toma un horizonte de planeación de siete periodos, y_1 sería 190. Con un horizonte de planeación de ocho periodos y_1 de nuevo se vuelve 400. Podría pensarse que este ciclo del valor de y_1 continuaría de manera indefinida. Ahora bien, resulta que no es el caso. Para horizontes de planeación con periodos $n \geq 21$, el valor de y_1 permanece fijo en 190.[3] Sin embargo, aun cuando existe una convergencia final de y_1, como en este ejemplo, la oscilación para los primeros 20 periodos podría ser problemática cuando se emplean horizontes de planeación progresivos.

Otro problema común que se presenta cuando se emplea el MRP es el *nerviosismo*. El término fue acuñado por Steele (1973), quien lo empleó para referirse a los cambios que ocurren en un esquema cuando el horizonte se desplaza hacia adelante un periodo. Algunas de las causas del nerviosismo incluyen cambios anticipados en el MPS como consecuencia de pronósticos actualizados, entregas tardías de materia prima, fallas de equipo clave, ausentismo de personal clave y variables impredecibles.

Se han realizado trabajos analíticos sobre el problema del nerviosismo. Carlson, Jucker y Kropp (1979 y 1983) emplean el término *nerviosismo* específicamente para señalar que un programa revisado requiere una preparación en un periodo en que el programa anterior no lo necesita. Dichos autores han propuesto una técnica interesante para reducir este tipo particular de nerviosismo: sea $(\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_N)$ el programa de producción existente y $(y_1, y_2, \ldots, y_N)$ un programa revisado basado en una nueva información de demanda. Supongamos que, además de los costos usuales de almacenamiento y preparación, existe un costo adicional de v si el nuevo programa $\mathbf{y}$ requiere una preparación en un periodo en que el antiguo programa $\hat{\mathbf{y}}$ no lo requería. Esto significa que existe un costo de instalación adicional en el caso del nuevo programa en aquellos periodos en los que no se necesitaba ninguna instalación en el antiguo programa. El método consiste en incrementar el costo de instalación de K a $K + v$ si $\hat{y}_k = 0$, antes de determinar el nuevo programa $\mathbf{y}$. Si se resuelve el problema con los costos de instalación modificados empleando cualquiera de los algoritmos para determinar tamaños de lote anteriormente analizados, ahora dará como resultado menos revisiones de preparación. Las ventajas de los diversos métodos para determinar tamaños de lote en este contexto se analizan en las dos referencias indicadas. El costo de revisión v refleja la importancia relativa del costo del nerviosismo.

---

[3] Se agradece a Lawrence Robinson de la Universidad de Cornell por señalar que esta anomalía no se extiende para todos los valores de n, como se había afirmado las ediciones anteriores, y en Carlson, Beckman y Kropp (1982).

Consideraciones adicionales

Aunque el MRP parecería la forma más lógica de programar la producción en procesos por lotes, el método básico posee serios inconvenientes, como hemos visto. Consideraremos otras dificultades.

Tiempos de entrega dependientes de los tamaños de lote

El cálculo del MRP supone que el tiempo de entrega de la producción de un nivel al siguiente constituye una constante fija independiente del tamaño del lote. En muchos contextos, esta afirmación resulta evidentemente ilógica. Se esperaría que el tiempo de entrega se incremente si aumenta el tamaño de lote. Si se incluye una dependencia entre el tiempo de entrega de la producción y el tamaño de la producción, el cálculo de la explosión parece extremadamente difícil.

MRP II: Planeación de recursos de manufactura

Según hemos notado, el MRP es un sistema de planeación de la producción cerrado que convierte un MPS en liberaciones de pedidos planeadas. La planeación de recursos de manufactura (MRP II) constituye una filosofía que pretende incorporar las demás actividades relevantes de la empresa en el proceso de planeación de la producción. En particular, las funciones financieras, contables y de mercadotecnia se vinculan a la función de las operaciones. Como ejemplo de la diferencia entre las perspectivas que ofrecen MRP y MRP II, revisemos el papel que el programa maestro juega en la producción. En el MRP, el MPS se considera como información de entrada. En el MRP II, el MPS se consideraría una parte del sistema y, como tal, sería también una variable de decisión. Por lo tanto, el gerente de control de la producción trabajaría con el gerente de mercadotecnia en determinar cuándo debería alterarse el programa de producción para incorporar revisiones en el pronóstico y nuevos compromisos de pedidos. En última instancia, todas las divisiones de la compañía trabajarían juntas para determinar un programa de producción compatible con el plan de negocios conjunto y la estrategia financiera a largo plazo de la compañía.

Otro aspecto importante de MRP II tiene que ver con incorporar la planeación de recursos de capacidad (CPR). Los aspectos de capacidad no se toman en cuenta explícitamente en el MRP. El MRP II es un ciclo de lazo cerrado donde el tamaño del lote y los programas del piso de producción asociados se comparan con las capacidades y se vuelven a calcular para cumplir con las restricciones de capacidad. Sin embargo, las cuestiones de capacidad continúan siendo un tema importante en los sistemas operativos MRP y MRP II.

Obviamente dicha aproximación global al problema de la programación de la producción resulta muy ambiciosa. Está por verse si dicha filosofía puede convertirse en un sistema factible en un entorno de operaciones particular.

Procesos de producción imperfectos

Una suposición implícita en el MRP es el hecho de que no se producen artículos defectuosos. Como los requerimientos de componentes y subensambles se calculan con el fin de satisfacer exactamente pronósticos de demanda de artículos finales, las pérdidas debidas a defectos pueden trastornar seriamente el equilibrio del plan de producción. En algunas industrias, como la de los semiconductores, las tasas de rendimiento pueden ser tan bajas como del 10 al 20 por ciento en el caso de productos nuevos. Siempre y cuando los rendimientos sean estables, la incorporación de pérdidas de rendimientos en el cálculo del MRP no resulta difícil. Se calculan las demandas netas y los tamaños de lote de la misma forma, y en la etapa final se dividen las cifras de requisiciones planeadas entre el rendimiento promedio. Por ejemplo, si un proceso en particular posee una tasa de rendimiento del 78 por ciento, se podrían multiplicar las requisiciones planeadas por $1/0.78 = 1.28$. El problema es mucho más complejo si los rendimientos son aleatorios y las varianzas demasiado grandes como para que se les ignore. Utilizar rendimientos medios resultaría en faltantes apreciables. Habría que diseñar una especie de modelo de repartidor de periódicos que equilibrara el costo de producir demasiados y muy pocos y determinar un factor de seguridad adecuado. A consecuencia de las dependencias de niveles su-

Aplicación instantánea

LA RAYMOND CORPORATION DESARROLLA LA MANUFACTURA DE CLASE MUNDIAL CON MRP II

La Raymond Corporation es una importante empresa manufacturera de transporte de materiales ubicada en Greene, Nueva York. Con el fin de lograr una amortización aproximada de 10 veces en inventarios reducidos y rendimientos, la organización llevó a cabo un proceso para poner en práctica el MRP II con una duración de dos años. El éxito de la empresa se basaba en diversos principios fundamentales. El primero tenía que ver con el hecho de que la alta gerencia tenía que comprometerse con el proceso de cambio. Por esa razón, el primer paso consistía en instruir a la alta gerencia utilizando un paquete educativo "enlatado" que poseía una firma consultora externa. El CEO facilitó la capacitación en un lugar adecuado para los vicepresidentes, quienes, a su vez, eran los responsables de la capacitación del equipo que conformaba la administración media. Con la alta gerencia a cargo, la administración media simplemente no podía pasar por alto el tema y esperar a que se olvidara. La capacitación continuó durante toda la fase del proyecto y en última instancia incluyó a los empleados de todo los niveles de la compañía.

Una vez que la capacitación avanzó suficiente para comenzar la realización del proyecto, se ideó una estrategia para hacerlo. El primer paso consistía en observar cuidadosamente los datos que se emplearían como entradas del sistema. Como resultado de este esfuerzo, la exactitud de la información del almacén pasó del 66 por ciento a cerca del 95 por ciento en aproximadamente 16 meses. Cada estado del inventario se medía en cuanto a exactitud y se corregía cuando era necesario. El establecimiento de normas y sistemas que proporcionan una información exacta puede constituir la mayor parte del beneficio de un esfuerzo como éste. Un resultado temprano fue que la exactitud de la información reducía la necesidad de supervisar las compras. Los informes de estado del inventario se empleaban para hacer una clasificación elemental, y se dedicaba más tiempo y energía al manejo de los artículos "A".

Con el fin de determinar la eficacia del nuevo sistema, el equipo de ejecución del proyecto se reuniría cada semana para repasar el avance de las medidas internas. Se hizo un intento para determinar la causa de los problemas o la falta de progreso sin señalar a nadie ni fincar culpas. Las medidas de desempeño se convirtieron en una parte importante del éxito del esfuerzo de ejecución. Las medidas exactas y actualizadas de desempeño constituyen la piedra angular de cualquier esfuerzo de cambio en los sistemas, pero se les debe aplicar sin amenazar a los trabajadores.

Por último, el autor recomienda que el paso final debería consistir en la compra de nuevos programas de cómputo. Si los datos subyacentes y los sistemas de medidas en determinado lugar no son adecuados, el programa, sin importar su sofisticación, no será de ninguna ayuda. Muchas compañías creen que la compra de un sistema MRP costoso es todo lo que se necesita para conseguir un inventario reducido, costos de materiales más bajos, mejoras en la entrega a tiempo, etc. Sin embargo, sin empleados a cargo, datos exactos y medidas de desempeño, el nuevo programa puede ser más un obstáculo que una ayuda.

Además de alcanzar el nivel clase A, ¿qué beneficios aprecia Raymond en este proceso? Sheldon (1994) mencionó la eliminación del tiempo extra, la eliminación de faltantes, una mejoría en entregas puntuales, reducciones en los tiempos y costos de preparación, reducciones en las tasas de reproceso y desperdicios, y costos reducidos de inventario y almacenamiento de material. Aunque resulta difícil creer que las mejoras fueron tan espectaculares como éstas, una ejecución del MRP II evidentemente valió la pena para Raymond Corporation.

Fuente: Esta descripción está basada en Sheldon (1994).

cesivos, esto sería un problema difícil para formular mediante un modelo matemático. La simulación por computadora Monte Carlo sería una buena alternativa. Hasta donde sabe el autor del libro, nadie ha intentado desarrollar un modelo matemático de rendimientos aleatorios en el contexto de los sistemas MRP. (Sin embargo, diversos investigadores analizan los rendimientos aleatorios. Por ejemplo, Nahmias y Moinzadeh, 1996, consideran un modelo matemático de rendimientos aleatorios en un problema de determinación de tamaños de lote de un solo nivel.)

Integridad de los datos

Un sistema MRP puede funcionar eficazmente sólo si los números que representan los niveles de existencias son exactos. Es fácil que se introduzcan datos incorrectos al sistema de programación. Esto puede ocurrir si un envío no se registra, o se anota incorrectamente en cierto nivel; los artículos que entran al inventario para ser reprocesados no se incluyen; las tasas de desperdicio son superiores que las anticipadas, etc. Con el fin de asegurar la integridad de los

datos empleados para determinar el tamaño y los tiempos de los lotes, puede requerirse implementar conteo físico del inventario a intervalos regulares. Una alternativa del conteo físico del inventario es una técnica conocida como *conteo de ciclos.* El conteo de ciclos simplemente significa verificar directamente los niveles disponibles de diversos inventarios que abarca el sistema MRP. Por ejemplo, ¿se indican las 45 unidades de la parte A557J en el registro actual del conteo real de este número de parte?

El conteo eficiente de ciclos puede conseguirse de muchas formas. Los almacenes pueden tener contenedores que sólo almacenan una cantidad fija o peso específico de artículos. Pueden emplearse sistemas de estantería codificada con el fin de identificar con mayor facilidad artículos con números de parte. Ciertas áreas podrían hacerse más accesibles sólo para personal específico. Los sistemas para el conteo de ciclos pueden basarse en el número o el peso. Además, debe considerarse en términos relativos un error en el nivel de existencias. Con base en la importancia del artículo, pueden aceptarse diferentes errores porcentuales. Deben aplicarse distintas tolerancias de error a los artículos contados por peso en comparación con los artículos contados por número. Si el MRP ha de tener un impacto positivo en el problema general de programación de la producción, los registros de inventario deben reflejar con exactitud el estado real del sistema.

Fijación de pedidos

En algunos sistemas complejos, se puede utilizar un solo componente en más de un artículo en el siguiente nivel superior del sistema. Por ejemplo, una compañía que produce muchos modelos de autos de juguete puede emplear el eje del mismo tamaño en cada uno de los automóviles. Los requerimientos brutos para los ejes serían la suma de los requerimientos generados por el MPS para cada modelo de automóvil. De ahí que cuando un componente se emplea en diversos artículos, el programa de requerimientos brutos para este componente proviene de diferentes fuentes. Si se presenta una escasez de este componente, para la compañía será útil poder identificar los artículos particulares en un nivel superior de la estructura de árbol que serían afectados. Para esto, el programa de requerimientos brutos divide los artículos que genera y cada requerimiento se *fija* con el número de parte de la fuente del requerimiento. La fijación añade una considerable complejidad a los requerimientos de almacenaje de información del sistema y sólo debería considerarse cuando la información adicional es importante en el proceso de toma de decisiones.

Problemas para la sección 7.5

30. Los sistemas MRP se han empleado con diversos grados de éxito. Describa bajo qué circunstancias debe el MRP ser exitoso y bajo qué circunstancias puede no serlo.
31. Comente las ventajas y desventajas de incluir una existencia de seguridad en los cálculos de tamaño de lote MRP. ¿Cree usted que un gerente de control de producción se mostraría renuente a construir una existencia de seguridad si se encuentra retrasado con respecto al programa?
32. ¿Por qué razón no resulta adecuado el método para determinar tamaños de lote con capacidad, analizado en la sección 7.5, para resolver el problema general de capacidad?
33. En la siguiente tabla se presentan las requisiciones planeadas (RP) para los tres componentes, A, B y C. Suponga que los rendimientos para estos componentes son, respectivamente, 84%, 92% y 70%. Suponiendo una programación lote por lote, ¿cómo se ajustan estas requisiciones planeadas para explicar el hecho de que los rendimientos son inferiores al 100%?

Semana	6	7	8	9	10	11	12	13	14	15	16	17
RP(A)				200	200	80	80	200	400	400	400	
RP(B)			100	100	40	40	100	200	200	200		
RP(C)	200	400	280	100	280	600	800	800	400			

34. Defina los términos *horizontes progresivos* y *nerviosismo del sistema* en el contexto de los sistemas MRP.

7.6 FUNDAMENTOS DE JIT

Justo a tiempo, producción esbelta y cero inventarios son todos nombres que indican esencialmente lo mismo: un sistema de flujo de materiales a través de la planta, el cual requiere un mínimo de existencias. Algunos han especulado que las raíces del sistema se remontan a la situación en el Japón de la posguerra. Devastadas por la guerra, las compañías japonesas carecían de efectivo y no se podían dar el lujo de invertir excesivamente en inventario. De esta manera, la producción esbelta nació de la necesidad. Sin embargo, conforme los autos japoneses comenzaron a ganar popularidad en los Estados Unidos, rápidamente se hizo evidente que eran muy superiores a los autos de fabricación americana y europea en términos de calidad, valor, eficiencia y confiabilidad. Ahora sabemos que el JIT y las iniciativas de calidad de los años cincuenta desempeñaron un papel importante en dicho logro.

Dos desarrollos resultaron decisivos en el éxito de este nuevo enfoque de la producción en masa: el sistema kanban y SMED (que significa sistema de cambio de dado en un minuto). Kanban es una palabra japonesa que significa *tarjeta* o *boleto*. Es un sistema de información manual creado y utilizado por Toyota para implementar el sistema justo a tiempo.

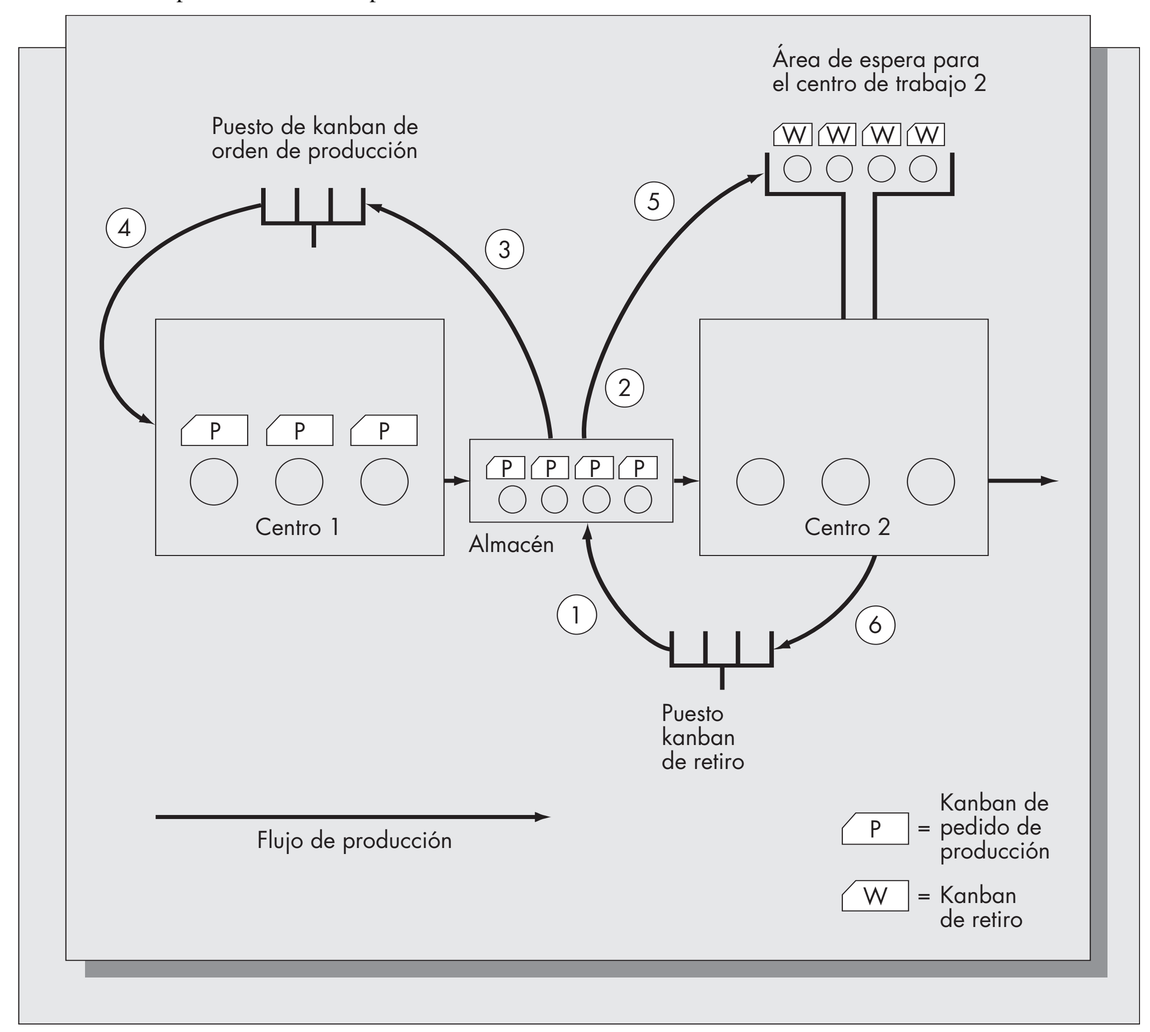
La mecánica del kanban

Existen varios tipos de etiquetas kanban, pero dos son los más comunes. Nos referimos a los kanbans de salida y los kanbans de orden de producción. Un kanban de salida es una solicitud de partes al centro de trabajo proveniente de un nivel superior del sistema. Un kanban de orden de producción es una señal para que un centro de trabajo produzca lotes adicionales. La forma en que estas dos etiquetas de kanban se emplean para controlar el flujo de producción se describe en la figura 7-8.

El proceso es el siguiente: las partes se producen en el centro de trabajo 1, se almacenan en una ubicación intermedia (conocida como almacén) y enseguida se transportan al centro de trabajo 2. Las partes se transportan en pequeños lotes representados por los círculos de la figura. La producción fluye de izquierda a derecha en el diagrama. Los pasos detallados del proceso son los siguientes (los números aparecen en los lugares adecuados de la figura 7-8):

1. Cuando el número de etiquetas en el kanban de salida alcanza un nivel determinado, un trabajador lleva estas etiquetas al almacén.
2. Si hay suficientes recipientes disponibles en el almacén, el trabador compara el número de parte en el kanban de orden de producción en el almacén con el número de parte en el kanban de salida.
3. Si los números de parte concuerdan, el trabajador retira los kanbans de orden de producción, y coloca los kanbans de salida en contenedores.
4. Cuando se ha acumulado una cantidad específica de kanbans de orden de producción, el centro de trabajo 1 procede con la producción.
5. El trabajador transporta partes recogidas en el almacén al centro de trabajo 2 y las coloca en un área de almacenamiento hasta que se les requiere en producción.
6. Cuando las partes entran a producción en el centro de trabajo 2, el trabajador retira los kanbans de salida y los coloca en el puesto para kanban de salida. (Obsérvese que los kanbans de orden de producción para el centro de trabajo 2 enseguida se anexan a las partes producidas en el centro de trabajo. Estas etiquetas de kanban no aparecen en la figura 7-8.)

FIGURA 7-8
Sistema kanban para dos centros de producción.



El número de etiquetas de kanban se calcula por adelantado en el sistema. Toyota emplea la siguiente fórmula (Monden, 1981b):

$$y = \frac{\overline{D}L + w}{a},$$

donde

y = Número de kanbans.

$\overline{D}$ = Demanda esperada por unidad de tiempo.

L = Tiempo de entrega (tiempo de procesamiento + tiempo de espera entre procesos + tiempo de transporte).

w = Política variable que especifica el nivel de la existencia de seguridad, generalmente de alrededor del 10 por ciento de $\overline{D}L$.

a = Capacidad del contenedor (por lo general no es mayor al 10 por ciento de la demanda diaria).

Esta fórmula implica que el máximo nivel del inventario está dado por $ay = \overline{D}L + w$. El valor ideal de w es cero. Sin embargo, resulta difícil equilibrar un sistema tan perfectamente que la existencia de seguridad se elimine por completo.

Como antes se mencionó, el sistema kanban es un sistema manual de información para llevar a cabo el sistema justo a tiempo. Los sistemas JIT también pueden ejecutarse de otras formas, que pueden resultar más eficientes que el método kanban. Se ahondará a este respecto más adelante en esta sección.

Cambio de dado en un minuto

Uno de los componentes fundamentales del éxito del sistema de producción de Toyota fue el concepto de cambio de dado en un minuto (SMED), propuesto por Shigeo Shingo. A Shingo generalmente se le atribuye la creación y ejecución de SMED en Toyota en 1970, sistema que se ha convertido en una parte importante del sistema general de producción de Toyota. La teoría básica expuesta en el capítulo 4 dice que los pequeños lotes serán óptimos sólo si los costos fijos son pequeños (conforme K disminuye en la fórmula CEP, también lo hace el valor de Q). El componente más significativo del costo de preparación de una nueva operación en una planta es el tiempo que se requiere para sustituir la maquinaria que llevará a cabo dicha operación, puesto que la línea de producción debe detenerse durante la operación de cambio. Esto requiere cambiar algunos juegos de herramientas y/o dados que se necesitan en el proceso; de ahí el nombre del término SMED. (Un dado es una herramienta empleada para dar forma o hacer una impresión en un objeto o material.)

Las operaciones de cambio de dados se llevan a cabo en las plantas de automóviles cuando se realiza el cambio de línea de producción de un modelo de auto a otro. Estas operaciones normalmente tardan cerca de cuatro horas. La premisa detrás de SMED consiste en que una porción significativa de la operación de cambio de dados puede realizarse mientras los dados anteriores aún se encuentran colocados y la línea continúa funcionando. De acuerdo con Shingo (1981), esto se logra dividiendo la operación de cambio de dados en dos componentes: un cambio interno de herramientas (IED, *internal exchange of die*) y un cambio externo de herramientas (OID, *outside exchange of die*). La operación OID se lleva a cabo mientras la línea funciona antes del cambio real. El objetivo es estructurar el cambio de dados de tal forma que haya tantos pasos como sea posible en la parte de OID de la operación.

Aunque esta idea suena simple, ha dado como resultado mejoras impresionantes en los ritmos de proceso y transferencia:

- En Toyota, una operación para realizar el cambio de dados y herramientas en un fabricante de pernos se redujo, en un año, de ocho horas a sólo 58 segundos.
- En Mitsubishi Heavy Industry, una operación de cambio de dados para una barrenadora se redujo de 24 horas a sólo 2 minutos y 40 segundos.
- En H. Weidmann Company de Suiza, el cambio de dados en una máquina para moldear plásticos se redujo de 2.5 horas a sólo 6 minutos, 35 segundos.
- En Federal-Mogul Company de Estados Unidos, el tiempo requerido para cambiar herramientas en una fresadora se redujo de 2 horas a 2 minutos.

Por supuesto, SMED no puede aplicarse en todos los contextos manufactureros. Incluso en contextos en los que puede aplicarse, los beneficios de la reducción en el cambio de dados se logran sólo si el proceso se integra a un sistema de control de fabricación general cuidadosamente diseñado y ejecutado.

Ventajas y desventajas de la filosofía justo a tiempo

Los defensores del sistema justo a tiempo creerían que los demás sistemas de planeación de la producción son ahora obsoletos. El celo con el que promueven el método recuerda el entusiasmo con que se anunció el MRP en la primera parte de la década de 1970. En ese tiempo, algunos afirmaron que el análisis clásico de inventarios ya no era válido. Sin embargo, cada nuevo método de producción debe verse como algo añadido, más que un remplazo, a lo que ya se

conoce. El sistema justo a tiempo puede ser una herramienta útil en las circunstancias correctas, pero está lejos de ser la panacea de los problemas que hoy enfrenta la industria.

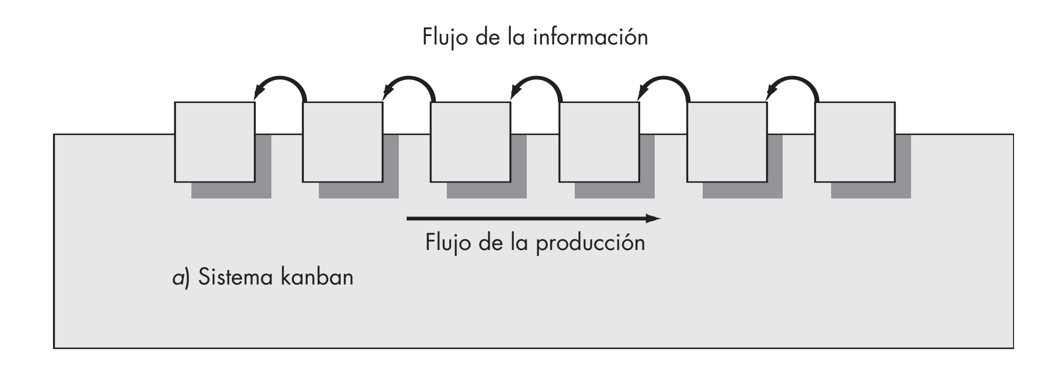
Los sistemas justo a tiempo y CEP no se excluyen. La reducción del tiempo de preparación y, como consecuencia, del costo de preparación, da como resultado tamaños de lote más pequeños. Los tamaños de lote más pequeños requieren mayor eficiencia y confiabilidad del proceso de producción, pero considerablemente menos inversión en materia prima, inventarios de productos en proceso e inventarios de bienes terminados.

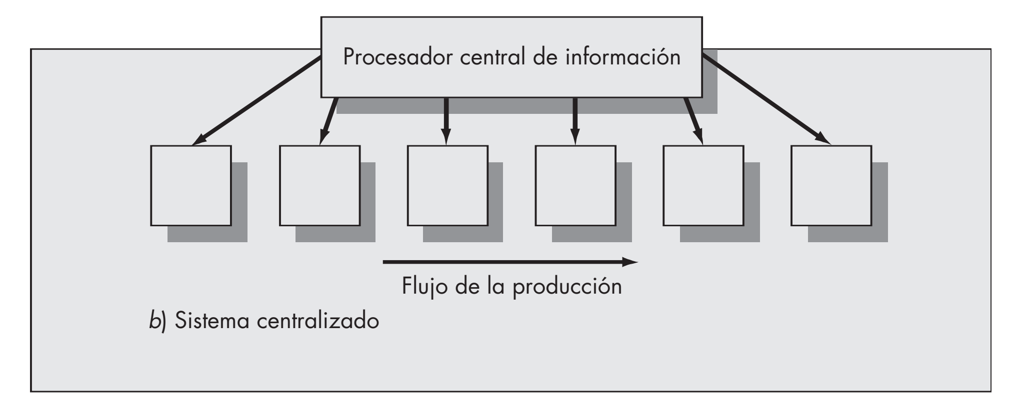
Kanban es un sistema manual de información empleado para apoyar el control de inventarios justo a tiempo. Los sistemas kanban y justo a tiempo no necesariamente están casados. En años futuros, esperamos ver sistemas justo a tiempo basados en tecnología de transferencia de información más actual y sofisticada. Un inconveniente de kanban es el tiempo que se requiere para transmitir nueva información a través del sistema. La figura 7-9 muestra un esquema de un proceso de producción en serie con seis niveles. Con el sistema kanban la dirección del flujo de la información es opuesta a la dirección del flujo de la producción. Consideremos las consecuencias de un cambio repentino en los requerimientos del nivel 6. Este cambio se transmite primero al nivel 5, enseguida al nivel 4 y así sucesivamente. Podría haber un lapso sustancial del instante cuando se presenta el cambio en el nivel 6 hasta que la información alcanza el nivel 1.

Un sistema de procesamiento de la información centralizado ayudará a remediar el problema. Si hay cambios repentinos en los requerimientos en un extremo del sistema, como resultado de cambios no planeados en la demanda o fallas en el equipo fundamental, estos cambios se transmitirán de forma instantánea a todo el sistema.

El MRP posee una ventaja importante sobre el kanban en este sentido. Uno de los aspectos fuertes de MRP es la capacidad para reaccionar a cambios pronosticados en el esquema de demanda. El sistema MRP vuelve a calcular cantidades de producción sobre la base de estos cambios y pone esta información simultáneamente a disposición de todos los niveles. El MRP

FIGURA 7-9
Sistema de información kanban en comparación con el sistema de información centralizado.





permite que la planeación tenga lugar en todos los niveles de una forma en que el sistema justo a tiempo y particularmente kanban no lo harían. Como Meal (1984) hace notar:

> La producción justo a tiempo funciona bien cuando el ritmo de producción en general es constante, aunque resulta insatisfactorio para comunicar cambios básicos en el ritmo de producción a etapas anteriores en el proceso… Por otra parte, al utilizar el enfoque HPP (*hiererchical production planning*, planeación de la producción jerárquica), los gerentes de planta no confían en sus señales de corto plazo para establecer sus ritmos de producción en las etapas iniciales.

El sistema justo a tiempo es más eficiente cuando el esquema de demanda es estable y predecible. Los cambios en la demanda pueden ser el resultado de causas predecibles, de fluctuaciones aleatorias, o de ambas. El MRP hace uso de pronósticos de cambios anticipados en la demanda y transmite esta información a todas las partes del sistema productivo. Sin embargo, ni el sistema MRP ni el sistema justo a tiempo están diseñados para protegerse de fluctuaciones aleatorias de la demanda. Ambos métodos podrían ser inestables frente a variaciones altas en la demanda.

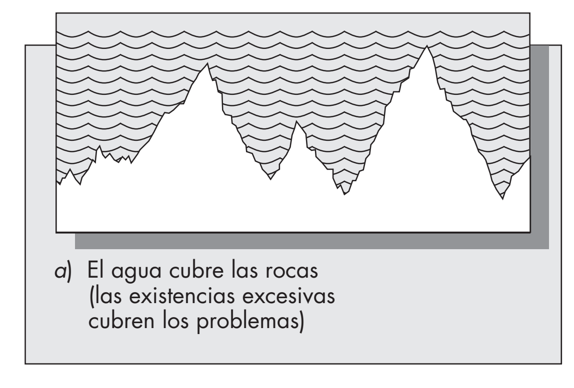
Otra desventaja potencial del sistema justo a tiempo es el tiempo muerto que puede resultar cuando se presentan descomposturas no programadas. Parte de la filosofía japonesa consiste en que los trabajadores deben estar familiarizados con más de una parte del proceso de producción. Si se presenta una falla, entonces la atención de los trabajadores puede enfocarse de inmediato en el problema. Sin embargo, si los trabajadores se encuentran familiarizados exclusivamente con su propia operación, habrá una cantidad significativa de tiempo muerto cuando se presente la falla. Esto concuerda con la curva de relación de la figura 8-4 del capítulo 8 relativo al control en el piso de producción y la programación en secuencia. Las existencias de seguridad entre operaciones sucesivas proporcionan un medio de hacer que los procesos de producción funcionen sin contratiempos. Sin embargo, los inventarios de seguridad también tienen sus desventajas. Pueden cubrir problemas subyacentes. Una analogía popular consiste en comparar un proceso de producción con un río y el nivel existencias con el nivel del agua en el río. Cuando el nivel del agua es alto, el agua cubre las rocas. Asimismo, cuando los niveles de existencias son altos, los problemas quedan cubiertos. Sin embargo, cuando el nivel del agua (las existencias) es bajo, las rocas (los problemas) se hacen evidentes (figura 7-10).

Como los artículos se desplazan a través del sistema en pequeños lotes, es factible la inspección al 100%. Desde esta perspectiva, el sistema justo a tiempo puede incorporarse con facilidad a una estrategia de control de calidad general. La administración de la calidad total (TQM), que se estudia en el capítulo 11, y el sistema JIT pueden funcionar juntos no solamente para reducir los costos de inventario, sino para generar importantes mejoras en la calidad de la producción.

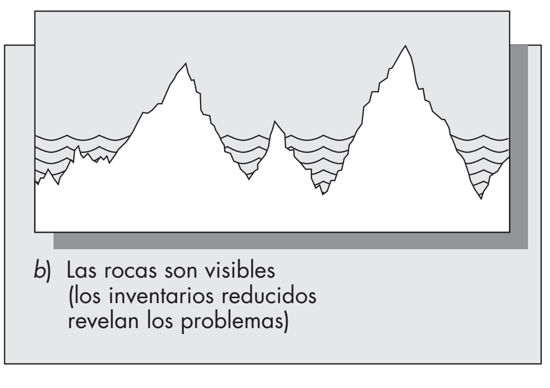
Como lo señalan los usuarios del sistema justo a tiempo, éste no es simplemente un sistema de control de existencias. Para que el sistema justo a tiempo funcione adecuadamente, debe coordinarse con el sistema de compras y con las estrategias de compras. Una queja relacionada con el sistema justo a tiempo tiene que ver con el hecho de que empuja la incertidumbre del

FIGURA 7-10

Analogía río-inventario, que ilustra las ventajas del sistema justo a tiempo.



a) El agua cubre las rocas (las existencias excesivas cubren los problemas)



b) Las rocas son visibles (los inventarios reducidos revelan los problemas)

TABLA 7-1 Resumen de las ventajas y desventajas de los sistemas justo a tiempo y Kanban

Características	Ventajas	Desventajas
Pequeños inventarios de productos en proceso	1. Reduce los costos de inventario. 2. Mejora la eficiencia de la producción. 3. Indica rápidamente problemas de calidad.	1. Puede dar como resultado un incremento en el tiempo muerto de trabajo. 2. Puede reducir el ritmo de producción.
Sistema de flujo de información kanban	1. Permite el rastreo eficiente de lotes. 2. Medios poco costosos de implementación del sistema justo a tiempo. 3. Permite un nivel predeterminado de inventario de producto en proceso presentando la cantidad de etiquetas de *kanban*.	1. Lento para reaccionar a los cambios en la demanda. 2. No toma en cuenta la información conocida sobre los esquemas de demanda futura.
Inventarios y compras coordinados	1. Reducción de inventarios. 2. Mejor coordinación de diferentes sistemas. 3. Mejores relaciones con los proveedores.	1. Menor oportunidad de abastecimiento múltiple. 2. Los proveedores deben reaccionar más rápidamente. 3. Se requiere mayor confiabilidad en los proveedores.

sistema y los inventarios grandes hacia el proveedor. No hay duda de que se requiere una gran flexibilidad por parte de los proveedores; éstos deben ser capaces de reaccionar con rapidez y proporcionar partes suficientemente confiables para evitar que el fabricante realice una inspección de todos los lotes que ingresan. Además, la localización de muchas fuentes de abastecimiento se torna difícil con dicho sistema. Es decir que la firma se puede ver forzada a tratar con un solo proveedor a fin de forjar la estrecha relación que exige el sistema. La localización de una fuente única presenta riesgos tanto para proveedores como para fabricantes. El fabricante enfrenta el riesgo de que el proveedor no pueda abastecerlo de partes cuando se requieren, y el proveedor enfrenta el riesgo de que el fabricante sufra reveses y disminuya la demanda.

La tabla 7-1 resume brevemente las ventajas y desventajas principales discutidas en esta sección.

Concluimos observando que evidentemente los sistemas JIT y MRP no son los únicos métodos para lograr el control de la manufactura. En su obra, Hopp y Spearman (1996) exploran un sistema de planeación de la producción basado en la fijación de un inventario de productos en proceso. Si la reducción de productos en proceso constituye un objetivo del JIT, entonces ¿por qué no diseñar un sistema que exija un nivel de producto en proceso deseado? Diseñaron justamente dicho sistema, lo denominaron CONWIP (trabajo en proceso constante, por la expresión en inglés, *CONstant Work-In Process*). En un sistema CONWIP, cada vez que se termina un artículo (que sale de la línea de producción), se inicia la producción de un nuevo artículo al principio de la línea. De esta manera el sistema WIP permanece fijo. Se puede ajustar el tamaño del WIP con base en aspectos de variabilidad y costo. CONWIP es similar en principio a JIT en el hecho de que ambos son sistemas *de tracción.* Sin embargo, en el caso de CONWIP, el flujo de información no procede de una etapa del proceso a una etapa anterior, sino del final del proceso al inicio. El método parece interesante, pero todavía necesita pasar la prueba de implementación en la realidad.

Implementación de JIT en Estados Unidos

Un ejemplo sorprendente de implementación exitosa del JIT en Estados Unidos es la Harley-Davidson Motorcycle Company. Hasta la primera mitad de la década de 1980, Harley-Davidson

era propiedad de la American Foundry Company (AMF). Harley, conocida como fabricante de motocicletas de cilindrada grande, encaró la fuerte competencia de los japoneses. Honda, por tradición un fabricante de motocicletas de pequeña cilindrada, comenzaba a hacer incursiones en el mercado de Harley. Parecía que las motocicletas se convertirían en otro producto de consumo producido exclusivamente por compañías no estadounidenses.

El primer paso en la recuperación de Harley fue que un grupo de empleados comprara la compañía a AMF. Inmediatamente después de la compra, la alta gerencia viajó a Japón para ver cómo funcionaban las fábricas de sus competidores. De acuerdo con Willis (1986),

> El *secreto* real, descubierto por los ejecutivos, no se encontraba en los sistemas de robótica o de alta tecnología, sino en la organización inteligente y eficaz de los empleados de la compañía y los sistemas de producción.

Como consecuencia de estas visitas, Harley-Davidson se reorganizó completamente. La estructura directiva tradicional se remplazó por un sistema donde cada empleado era propietario de su propia área en la línea. Una gran proporción de personal se eliminó, lo cual dio como resultado un diagrama de organización mucho menos elaborado y similar al de las compañías japonesas. De esta manera, los problemas no podían ser enterrados por la burocracia. La firma también instituyó un programa de círculos de calidad con la participación activa del personal de la línea (véase el análisis de los círculos de calidad en el capítulo 11).

La empresa invirtió en el diseño y la manufactura asistidos por computadora (CAD/CAM) y equipo de robótica. La distribución física de la planta se reestructuró en células de trabajo utilizando una distribución de tecnología de grupo (véase capítulo 10). Las motocicletas ya no se fabricaban en una línea de ensamblaje tradicional.

Un cambio fundamental fue la forma en la que el material se desplazaba por la planta. Harley adoptó el término *MAN* (*materials as needed*, materiales según se les necesite) para describir este sistema, aunque éste es evidentemente otro nombre de JIT. Para hacer que el JIT funcionara, instituyó un programa para reducir el número de preparaciones diarias. Antes de la reestructuración, producía de 13 a 14 diferentes modelos cada día. Con este nuevo sistema, los pronósticos mejorados permitían implementar ciclos repetitivos de dos horas.

Con el fin de hacer que JIT funcionara, Harley necesitaba un compromiso de sus proveedores así como de sus propios empleados. Para estimular a sus vendedores a comprar en el programa, Harley ofreció sesiones de entrenamiento en los principios del JIT. Estos programas se hicieron tan populares que Harley comenzó a ofrecerlos a compañías que no eran ni proveedores ni clientes. ¿Cuál es el resultado de esta iniciativa? Hoy las motocicletas Harley-Davidson están un 99 por ciento libres de defectos, en comparación de apenas el 50 por ciento en 1982. Además de lograr sorprendentes mejoras en la calidad del producto, Harley desarrolló nuevos programas de beneficios para los empleados y amplió su línea de producción. Todo esto se logró con una reducción simultánea de costos.

Harley-Davidson no es el único ejemplo de implementación exitosa del JIT en Estados Unidos. Jacobson y Hillkrik (1986) describen el cambio total de Xerox, que fue sitiada por la competencia de Kodak y varias firmas japonesas (especialmente Canon) en la primera mitad de la década de 1980. Xerox se enfrascó en un programa de autoexamen basado en un *benchmarking* competitivo y llevó a cabo una reestructuración completa de su negocio. La implementación de los sistemas JIT desempeñó un papel importante en este cambio. Por ejemplo, redujo el número de proveedores de la compañía de 5 000 a sólo 400, lo cual incluía una combinación de recursos del extranjero, como proveedores de bajo costo y recursos cercanos, con el fin de que se hicieran entregas a tiempo cuando se requiriera. Estos sistemas permitieron reducir los inventarios de partes de copiadoras en $240 millones.

La industria automotriz norteamericana ha hecho movimientos importantes hacia la adopción del JIT, pero aún queda un largo camino por recorrer. En Estados Unidos existen diferentes plantas que han adoptado la filosofía JIT. Dos empresas de General Motors son la planta de NUMMI, al norte de California, una empresa conjunta con Toyota, y la planta Saturn en Tennessee.

Irónicamente, Ford, que fue pionera de las líneas de ensamble para la fabricación en serie de automóviles, ha sido la firma norteamericana más exitosa al adoptar las filosofías de fabricación japonesas. La planta de ensamble de Ford cerca de Atlanta, Georgia, en donde se arman el Ford Taurus y el Mercury Sable, es impresionante en lo que se refiere a la mayoría de sus medidas de productividad y calidad, aun cuando se les compare con los japoneses. Cuando GM intentó aislar los factores responsables de la productividad significativamente mejor de Ford en esta planta, comparada con la planta de Fairfax de GM, que fabrica el Pontiac Grand Prix, se encontró que el 41 por ciento de la diferencia podía atribuirse a la facilidad de fabricación de los diseños automotrices y el 48 por ciento a las prácticas de manufactura (Womack y otros, 1990, p. 96).

Aunque las compañías de Estados Unidos han hecho progresos significativos, la conversión a la "producción esbelta" no ha resultado tan rápida como se hubiera esperado. Esto se debe en parte al entorno poco favorable para el JIT. Por ejemplo, en la industria automotriz de Estados Unidos, las plantas de ensamble y los proveedores con frecuencia se encuentran separados por grandes distancias. Las firmas japonesas pueden contar con entregas diarias, incluso en horas, en virtud de que los proveedores se localizan cerca de las plantas de ensamble. Asimismo, el JIT funciona mejor cuando la demanda de productos es relativamente estable. Las ventas de automóviles en Estados Unidos han resultado más cíclicas que en Japón (Womack y otros, 1990, p. 27), lo cual ha tenido como consecuencia que el sistema JIT sea más difícil de implementar ahí. Por consiguiente, la industria automotriz de Estados Unidos enfrenta un entorno menos favorable para la implementación exitosa del JIT. Aun con estas dificultades, tanto los fabricantes de Estados Unidos como los europeos tendrán que hacer la transición a los métodos de producción esbelta si esperan seguir siendo competitivos con respecto a los japoneses.

Problemas para la sección 7.6

35. Compare los conceptos de empuje y tracción y describa cómo se relacionan con el sistema justo a tiempo.
36. ¿Cuál es la diferencia entre un sistema justo a tiempo y un sistema kanban? ¿Puede un sistema justo a tiempo implementarse sin kanban?
37. Un fabricante regional de lámparas de mesa planea emplear un sistema manual de flujo de información kanban. En promedio, la firma produce 1 200 lámparas mensuales. El tiempo de entrega de la producción es de 18 días, y se planea contar con un 15 por ciento de *stock* de seguridad. Supongamos que el mes tiene 20 días laborables.

    a) Si cada contenedor almacena 15 lámparas, ¿cuál será el total de etiquetas kanban que se requieren? (Utilice la fórmula de esta sección.)

    b) ¿Cuál es el máximo inventario de productos en proceso que la compañía puede esperar tener con este sistema?

    c) Suponga que a la compañía le cuesta $30 producir cada lámpara. Si los costos por mantener inventarios se basan en una tasa de interés anual del 20%, ¿en qué costo anual de inventarios incurre la compañía para los productos en proceso? (Puede revisar el análisis de los costos por mantener inventarios en el capítulo 4.)

38. ¿Qué es el SMED? ¿De qué forma se pueden reducir las operaciones de cambio de dados con pedidos de cierta magnitud? ¿Por qué tiene ventajas hacerlo?
39. Explique la forma en que funciona el sistema kanban de dos etiquetas.
40. Discuta las ventajas y desventajas de cada una de las siguientes características del sistema justo a tiempo:

    a) Tamaños pequeños de lote.

    b) Control integrado de compras e inventarios.

    c) Sistema de información kanban.

7.7 COMPARACIÓN ENTRE EL MRP Y EL JIT

MRP y JIT son fundamentalmente sistemas diferentes de control de manufactura. Como lo mencionamos antes, MRP es un sistema *de empuje* y JIT es un sistema *de tracción.* JIT es un sistema reactivo. Si se presenta un problema y la línea se detiene, JIT reacciona de inmediato, porque las órdenes de material nuevo se interrumpen. De esta manera, se podría decir que JIT reacciona ante incertidumbres y MRP no lo hace. Sin embargo, evidentemente JIT no funcionará bien cuando se sabe que las demandas variarán considerablemente con el tiempo. MRP construye la información en la estructura de planeación, mientras que JIT no lo hace.

En la mayoría de los entornos de manufactura, la implementación de un sistema puro JIT simplemente no es factible. Los proveedores quizás no se encuentren localizados suficientemente cerca para que las entradas se entreguen de acuerdo con un programa rígido. Las demandas de productos pueden variar considerablemente, lo cual hace poco práctico no tomar en cuenta esta información en el proceso de planeación. Puede resultar difícil desplazar productos en pequeños lotes. La implementación de SMED puede no ser posible en algunos entornos. Cuando los costos de preparación son muy altos, tiene sentido desde el punto de vista económico producir grandes lotes y almacenar artículos en lugar de cambiar los procesos de producción con frecuencia. Sin embargo, habiendo dicho esto, se pueden conseguir reducciones importantes de existencias WIP en la gran mayoría de las plantas de manufactura tradicionales. Las plantas más eficientes funcionan mejor.

El enorme éxito de Toyota en la reducción de costos de inventario mientras fabrica productos de alta calidad con grandes tasas de producción, ha formado la base de gran parte de los sistemas basados en el JIT. Sin embargo, no resulta claro si el JIT es el principal responsable del éxito de Toyota. La forma actual de hacer negocios de Toyota difiere de la de los fabricantes de automóviles norteamericanos en varias dimensiones. ¿Es su éxito un resultado inmediato de los métodos JIT, o puede atribuirse a otros factores que podrían ser más difíciles de emular? Con el fin de determinar bajo qué circunstancias el sistema JIT podría ser más ventajoso, Krajewski *et al.* (1987) idearon un simulador a gran escala para comparar los entornos de manufactura de JIT, MRP y ROP (punto de reorden). La comparación incluía 36 factores directos agrupados en ocho categorías principales, en las que cada factor variaba de uno a cinco niveles. A continuación, se enlistan las ocho categorías consideradas y un breve resumen de los factores incluidos en ellas:

1. *Influencia del cliente.* Error de pronóstico de la demanda.
2. *Influencia del proveedor.* Confiabilidad del proveedor (tamaño del pedido comparado con el tamaño de pedido solicitado, y tiempo de recepción comparado con el tiempo solicitado).
3. *Mecanismos de amortiguamiento.* Amortiguamiento para la capacidad y la existencia de seguridad, así como tiempos de entrega de seguridad.
4. *Estructura del producto.* Consideración de estructuras de pirámide (pocos artículos finales) y de pirámide invertida (muchos artículos finales).
5. *Diseño de instalaciones.* Patrones de rutas, énfasis en las máquinas y longitud de las rutas.
6. *Proceso.* Tasas de desperdicio, fallas en el equipo, flexibilidad del trabajador y desbalanceos de la capacidad.
7. *Inventarios.* Exactitud en los informes, reglas de determinación de tamaños de lote y tiempos de instalación.
8. *Otros factores.* Número de artículos, número de estaciones de trabajo y otros factores.

Los resultados obtenidos fueron sumamente interesantes. El JIT trabajó bien sólo en entornos de manufactura favorables. Esto implica poca o ninguna variabilidad, proveedores confiables y tiempos breves de preparación de la producción. El JIT mostró un desempeño pobre cuando uno o más de estos factores resultaron desfavorables. De hecho, en un entorno favorable, tanto ROP como MRP también dieron buenos resultados. Esto sugiere que el beneficio fundamental proviene de la creación de un entorno de manufactura favorable opuesto a una simple implementación del sistema JIT. Quizás se hubiera obtenido mayor beneficio de haberse evaluado cuidadosamente (cuando fuera posible) y se hubieran rectificado los problemas

fundamentales que afectaban la manufactura, en lugar de implementar a ciegas un nuevo sistema de planeación de la producción. Por otra parte, quizás es en el *proceso* de implementación del sistema JIT del que se obtiene el mayor beneficio.

Los nuevos enfoques de control de la manufactura se suelen etiquetar rápidamente con acrónimos de tres letras y se hace proselitismo con el celo de una religión. El JIT no es la excepción. En el mejor de los casos, JIT es un conjunto de métodos que deben implementarse sobre una base de mejoramiento continuo para reducir existencias en cada nivel de la cadena de abastecimiento. En el peor de los casos JIT constituye una idea romántica que, cuando se aplica a ciegas, puede ser muy peligrosa para la moral del trabajador, las relaciones con los proveedores y, en última instancia, para los resultados. Los ahorros en inventario a menudo pueden ser una ilusión, como lo demostraron Chhikara y Weiss (1995). Expresaron, en tres estudios de casos reales, que si las reducciones de existencias no se vinculan a sistemas contables y flujos de efectivo, no pueden traducirse en una reducción de costos de inventario. El JIT no tiene sentido en sí mismo, se le debe integrar cuidadosamente en toda la cadena de manufactura y abastecimiento para poder aprovechar sus beneficios.

Zipkin (1991) presenta un artículo muy profundo sobre la "revolución del JIT". Su tesis principal consiste en que el JIT es capaz de procurar beneficios reales, pero si el pragmatismo no se distingue del romanticismo, las consecuencias pueden ser desastrosas. Cita el ejemplo del gerente de una planta de ensamble de computadoras que ordenó reducir su inventario WIP a cero en seis meses. Ése fue el resultado, aparentemente, de las experiencias del CFO, tras asistir a un seminario de JIT que lo inspiró a promover reducciones masivas de existencias.

Por último, como lo señala Karmarkar (1989), la cuestión no es hacer una elección entre MRP y JIT, sino hacer el mejor uso de ambas técnicas. JIT reacciona muy lentamente a cambios repentinos en la demanda, mientras que MRP incorpora pronósticos de demanda en el plan. ¿Significa esto que Toyota, famosa por su sistema JIT, ignora los pronósticos de demanda en su planeación y control de manufactura? Es muy poco probable. La comprensión de lo que las diversas metodologías ofrecen, así como sus limitaciones, conduce a un sistema de control y planeación de manufactura bien diseñado y eficiente. Las mejoras deben ser incrementales, no revolucionarias.

7.8 SOFTWARE DE MRP: UN GRAN NEGOCIO

El software de control y planeación de la manufactura es un gran negocio y continúa creciendo rápidamente. El software de planeación de requerimientos de materiales salió a la luz en la década de 1970. Estos sistemas fueron diseñados para computadoras *mainframe*, que con frecuencia mezclan hardware y software en un solo paquete. Sin embargo, pronto se hizo evidente que los programas de planeación de manufactura independientes se estaban quedando cortos porque no lograban los vínculos necesarios con otras partes de las operaciones de la compañía. La segunda generación de software de MRP (MRP II o planeación de recursos de manufactura) intentó incorporar los *recursos* que tendían un vínculo con el plan de manufactura, incluyendo proyección de ventas, personal, capacidad de la planta y capacidad de distribución.

¿Cuán grande es el mercado? LaPlante (1994) informó que los ingresos que se habían proyectado para dichos sistemas (tanto MRP como MRP II) para 1994 fueron de $2.8 mil millones en Estados Unidos, lo cual representaba el 40 por ciento de crecimiento sobre 1993. De hecho la tasa de ventas continuó en ascenso en 1995 sin señales de baja. Blood (1994) informó que más de 70 000 sistemas de este tipo se habían instalado en Estados Unidos.

Con este rápido crecimiento y decenas de miles de usuarios solamente en Estados Unidos, podría suponerse que los clientes están felices con este software. Sin embargo, esto parece estar lejos de la verdad. Blood (1994) afirma que dos tercios de los usuarios de MRP no se encuentran satisfechos con el producto. De acuerdo con George Stalk, hijo: "Los esfuerzos de automatización de la fábrica que reducen costos y mejoran la eficiencia de forma ascendente capturan sólo el fruto que pende más bajo. La mayoría de las compañías se encuentra corrigiendo

las ineficiencias provocadas por la falta de atención. Están fallando rotundamente en obtener realmente grandes rendimientos" (citado en LaPlante, 1994).

No hay garantía siquiera de que "se cortará el fruto más bajo". El nuevo software puede empeorar las cosas en lugar de mejorarlas. Los problemas de implementación provienen de dos fuentes. Los usuarios esperan que un día después de que un nuevo sistema se ha instalado, las existencias desciendan, las entregas mejoren y todo marche sin contratiempos. Compárese esta actitud con la preparación cuidadosa de Raymond Corporation analizada antes en la Aplicación instantánea. A Raymond le llevó muchos años de entrenamiento instalar el nuevo software.

El otro problema surge de las limitaciones en los sistemas mismos. La mayoría de los sistemas MRP y MRP II en el mercado todavía emplean computadoras *mainframe* y códigos fuente complicados que no incluyen vínculos con las bases de datos de mercadotecnia o contabilidad. Mientras muchos de éstos tuvieron éxito en el pasado (como el sistema MANMAN de Ask Sistemas de Mountain View, California), el crecimiento real en el mercado proviene de otra fuente. Muchas compañías ofrecen configuraciones más modernas basadas en sistemas cliente-servidor. En una estructura cliente-servidor, las máquinas satélite (normalmente computadoras personales) se vinculan mediante un servidor de tal manera que se pueda tener acceso a bases de datos comunes. La interfaz de la PC facilita la introducción y la recuperación de datos en ambientes familiares, como la hoja de cálculo. También permite a diversas partes de la organización obtener información sobre el estado de los pedidos importantes.

Entre las compañías que ofrecen sistemas abiertos cliente-servidor están Avalon, Baan, qad.inc, Oracle, SAP, Symix y Triton. Aunque todas estas empresas han tenido cierto éxito, la que parece estar incursionando más rápidamente en el mercado de la manufactura es SAP. Esta empresa alemana que vende computadoras, programas y servicios de consultoría en todo el mundo, proyectó ganancias de $9 300 millones en 1995. Los sistemas SAP enlazan la fábrica, la bodega, las oficinas de ventas, de contabilidad y el centro corporativo en un sistema de información comprensible. La empresa afirma que tiene 4 500 clientes en todo el mundo.

Con SAP, el objetivo es pedir que los datos se introduzcan en el sistema sólo una vez. Si se vinculan todas las áreas funcionales de la compañía, estos datos quedan disponibles donde y cuando se les necesita. SAP (y sus competidores) son, por supuesto más que sólo sistemas de planeación de manufactura. Son sistemas de información de dirección integrada con múltiples funciones de planeación. Estos sistemas no son baratos. Un sistema de gran escala, que incluya hardware y asesoría de instalación, podría rondar alrededor de las decenas de millones de dólares. Rara vez podría adquirirse un producto SAP por menos de $100 000. Por esa razón, continuará existiendo un mercado de alternativas menos costosas.

Enseguida se presentan varios estudios de compañías que han tenido experiencias positivas con la instalación de la nueva generación de software de manufactura cliente-servidor. (Los primeros cuatro se encuentran descritos en LaPlante, 1994, y el último en Lieber, 1995.)

1. Northrup Grumann de Long Beach, California, dio de baja un sistema basado en una computadora *mainframe* que le estaba costando a la firma $20 millones anuales en mantenimiento y licencias, y lo remplazó con un sistema cliente-servidor de Oracle. Aun cuando el sistema Oracle costó $14 millones, los costos generales se redujeron de inmediato y el tiempo de entrega se redujo de 18 meses a sólo 7 meses.

2. Coors Ceramics de Golden, Colorado, padeció los males de muchas empresas manufactureras. Los números impresos por las computadoras no eran confiables, el estado de los pedidos no se podía determinar y los tiempos de entrega eran demasiado prolongados. Coors sustituyó un viejo sistema que corría en una DEC PDP 11 con un sistema cliente-servidor que vinculaba la manufactura, el servicio en campaña y las finanzas. El resultado fue que los ciclos de producto, que variaban de 10 a 14 semanas, se redujeron de 6 a 10 semanas y las entregas a tiempo mejoraron cerca del 95 por ciento.

3. Phoenix Designs es una empresa manufacturera que fabrica por pedido sistemas de oficina y muebles. Phoenix implementó un sistema de información Symix en un servidor IBM RS/6000 conectado a un programa de diseño asistido por computadora (CAD) localizado directamente en las tiendas. El sistema permitía a los comerciantes colocar un sistema Phoenix en una pantalla de computadora utilizando un indicador y un ratón. Empleaba una aplicación basada en una regla con principios de inteligencia artificial. El software fue utilizado por 236 comerciantes y se encuentra conectado al sistema de información de manufactura de la firma. Antes de la instalación el sistema, se requirieron de cuatro a seis semanas a partir del tiempo en que se hacía un pedido hasta que el producto se entregaba al cliente. Con el nuevo sistema, un sistema por pedido puede entregarse en dos semanas.

4. Red Devil, un productor de selladores, impermeabilizantes y herramientas con $50 millones en activos, se dio cuenta de que los empleados habían adquirido una desconfianza muy arraigada en el sistema de planeación de manufactura, y por una buena razón: los números estaban definitivamente incorrectos. Red Devil corrió un software Avalon en un servidor NCR Unix y consiguió ahorros inmediatos en costos sistemas de información de manufactura, así como un considerable ahorro de $2 millones en costos de inventario. El servicio al cliente pasó de una tasa del 95 por ciento en cinco días a una tasa del 98 por ciento en dos días.

5. El gigante de los alimentos Borden descubrió que la instalación del software y hardware SAP resolvía diversos problemas. En 1994, la firma instaló el SAP para el procesamiento de pedidos. Con un nuevo sistema en el lugar de trabajo, los empleados de cualquier lugar en el país podían rastrear el estado de un pedido mientras el cliente esperaba. En el pasado, a menudo llevaba un día entero localizar esta información.

Conforme los sistemas de información para la planeación de manufactura se tornen más complejos y fáciles de usar, las compañías podrán conseguir la eficiencia que no lograban en el pasado. Con la información disponible, estos sistemas algún día podrían incorporar el tipo de modelos matemáticos discutidos en este libro con el fin de conseguir todavía más eficacia.

7.9 NOTAS HISTÓRICAS

El término específico *MRP* es relativamente nuevo, aunque el concepto de planeación de materiales basado en la demanda predicha y las estructuras del producto parece que han estado presentes algún tiempo ya. De hecho, existe una referencia de este enfoque hecha por Alfred Sloan (1964), quien se refiere a un cálculo estrictamente técnico llevado a cabo por los analistas de General Motors en 1921 respecto a la cantidad de materiales que se necesitaban para la producción de determinada cantidad de automóviles. El término *explosión BOM* (lista de materiales, *bill of materials*) se empleó comúnmente para describir lo que ahora se denomina MRP.

Las obras de Orlicky (1975) y New (1974) contribuyeron a legitimar el MRP como técnica válida e identificable, aunque el término parece que ha estado presente desde mediados de la década de 1960. Además, a los conocidos profesionales George Plossl y Oliver Wight también se les debe dar crédito por popularizar el método (véase, por ejemplo, Plossl y Wight, 1971). En Anderson, Schroeder, Tupy y White (1982) los autores afirman que los primeros sistemas computadorizados MRP fueron implementados cerca de 1970. El número de sistemas instalados se ha incrementado desde aquel tiempo a un ritmo exponencial.

Es interesante que mucho del trabajo relacionado con los métodos de determinación óptima y no óptima de tamaños de lote antecede al reconocimiento formal del MRP. El papel original en esta área se debe a Wagner y Whitin (1958), quienes reconocieron por vez primera la naturaleza óptima de una política de requerimientos exactos para sistemas de control de inventarios de revisión periódica con demanda de variación de tiempos, y desarrollaron el algoritmo de programación dinámica descrito en este capítulo. La heurística de Silver-Meal, cuando las oportunidades de abastecimiento se encuentran al principio de los periodos, apareció en Silver y Meal (1973). A DeMatteis (1968) normalmente se le da el crédito por el enfoque de equilibrio de periodos parciales. Sin embargo, el artículo de Gorham (1968) se refiere tanto al equilibrio de periodos parciales (denominado *enfoque de costo mínimo total*) como a los métodos de costo mínimo unitario, lo cual sugiere que estos métodos se conocían bien en esa época. Es

probable que ambos métodos fueran desarrollados por profesionales antes de 1968, aunque no aparezcan en la literatura.

El algoritmo de cambio de lotes para el problema de capacidad bosquejado en la sección 7.4 es muy similar al que ideó Karni (1981). Dixon y Silver también exploraron ideas similares. La determinación de tamaños de lote no siempre fue usada por los profesionales en los sistemas operativos MRP, porque los errores pequeños a un nivel alto de la estructura del producto se propagan en forma de errores grandes en los niveles inferiores. Además, los algoritmos para determinar tamaños de lote requieren la aproximación de los costos de preparación y almacenamiento, y más cálculos que una simple política de lote por lote.

Las referencias históricas al JIT se encuentran incluidas en el capítulo. No está claro quién acuñó el término *JIT,* pero el concepto se deriva evidentemente del sistema kanban de Toyota. SMED (cambio de dados en un minuto) ha desempeñado un papel importante en el éxito de los métodos japoneses de producción ajustada. Por lo general, el crédito por este sistema se le da a Shigeo Shingo.

7.10 Resumen

La *planeación de requerimientos de materiales* (MRP) es un conjunto de procedimientos que convierten pronósticos de demanda de un producto manufacturado en un programa de requerimientos para los componentes, los ensambles y la materia prima que conlleva dicho producto. Un concepto relacionado es el de *programa maestro de producción* (MPS), que constituye una especificación de las necesidades previstas del artículo final por periodo de tiempo. El *cálculo de la explosión* es un conjunto de reglas y procedimientos para convertir el MPS en requerimientos para los niveles inferiores. La información que se requiere para llevar a cabo el cálculo de la explosión se encuentra contenida en el *diagrama de estructura del producto* y en la *lista indexada de materiales.* Las dos piezas de información clave contenidas en el diagrama de estructura del producto son los tiempos de entrega necesarios para producir el componente específico y el multiplicador que da el número de unidades del componente que se requiere para producir un artículo en el siguiente nivel superior de la estructura del producto.

Muchos sistemas MRP se basan en un programa de producción *lote por lote.* Es decir que el número de unidades de un componente producidas en un periodo es el mismo que los requerimientos para dicho componente en ese periodo. Sin embargo, si los costos de almacenamiento y preparación pueden calcularse con exactitud, es posible definir otras reglas para determinar tamaños de lote más económicos. El procedimiento óptimo para determinar el tamaño de lote es el algoritmo de Wagner-Within. Sin embargo, el método rara vez se utiliza en la práctica, principalmente debido a la relativa complejidad de los cálculos que se requieren (aunque dichos cálculos se realizan en poco tiempo en una computadora). Analizamos tres métodos heurísticos que requieren menos cálculos que el algoritmo de Wagner-Within, aunque ninguno de estos métodos necesariamente dará como resultado la solución óptima. Estos son el la *heurística de Silver-Meal,* la *heurística de costo unitario mínimo* y el *balanceo de periodos parciales.*

También estudiamos el problema de la determinación dinámica de tamaños de lote cuando existen restricciones de capacidad. Una de las limitaciones del MRP es el hecho de que no se tomen en cuenta las capacidades. Esto es particularmente importante si la determinación del tamaño del lote se incorpora en el sistema. La determinación de soluciones óptimas para un sistema de inventarios con capacidad restringida sujeto a una demanda que varía con el tiempo constituye un problema extremadamente complicado. (Para quien esté familiarizado con el término, el problema se dice que el sistema es NP completo, que es una referencia al nivel de dificultad.) Se expuso un método heurístico directo para obtener una solución del problema de la determinación de tamaños de lote con capacidad. Sin embargo, la incorporación de dicho método en el sistema MRP *no* resolverá por sí mismo el problema completo del MRP con capacidad, ya que si incluso un programa particular para determinar tamaños de lote resulta factible en cierto nivel, no existe garantía de que dará como resultado un programa factible de requerimientos en un nivel inferior.

Las soluciones realmente óptimas para la determinación de tamaños de lote en un sistema MRP requerirían que el problema se formulara en forma de un programa íntegro con el fin de determinar simultáneamente las decisiones óptimas para todos los niveles. En el caso de sistemas de ensamble reales, que pueden llegar a tener de 10 a 15 niveles de profundidad, esto daría como

resultado un enorme problema de programación matemática. En vista de la gran cantidad de temas concernientes a la implementación del MRP, el beneficio marginal que se podría obtener de la optimización de niveles múltiples probablemente no justificaría el esfuerzo implicado.

El *nerviosismo del sistema* constituye un problema que surge cuando se instala un sistema MRP. El término se refiere a los cambios anticipados en un programa que resultan cuando el horizonte de planeación se adelanta un periodo. Otra dificultad consiste en el hecho de que en muchas circunstancias, los tiempos de entrega de la producción dependen de los tamaños de lote: el MRP supone que los tiempos de entrega de la producción se encuentran fijos. Otro problema más es el hecho de que los rendimientos en diversos niveles del proceso quizás no sean perfectos. Si es posible calcular por adelantado las tasas de rendimiento, en los cálculos, dichas tasas pueden descomponerse en factores de manera directa. Sin embargo, en muchas industrias, las tasas de rendimiento pueden resultar difíciles de calcular por adelantado.

La *planeación de recursos de manufactura* (MRP II) intenta abordar algunos problemas de implementación de MRP integrando las funciones financieras, contables y mercadotécnicas en la función de planeación de la producción.

La industria de software de MRP continúa creciendo a un ritmo acelerado. Las ventas alrededor del mundo de sistemas de planeación de la manufactura ascienden a decenas de miles de dólares. Los sistemas tradicionales MRP se basaban en computadoras *mainframe* y operaban sin vínculos con otras bases de datos dentro de la compañía. La generación más nueva de software de manufactura conecta todas las partes activas de la compañía y se basa en sistemas cliente-servidor en lugar de computadoras *mainframe*. Aunque se ha hablado mucho de que el JIT sustituirá al MRP, la evidencia apunta al hecho de que la industria del software de MRP continúa en expansión. Los sistemas de planeación de manufactura tipo MRP desempeñan un papel importante en los métodos gerenciales de la fábrica.

En este capítulo también se analizó el sistema justo a tiempo. Es un sistema de control de existencias cuyo objetivo es reducir los productos en proceso a un mínimo. Los conceptos se basan en los sistemas de control de la producción empleados por Toyota en Japón en la década de 1970. JIT es un sistema *de tracción,* mientras que MRP es un sistema *de empuje.* Las partes se transfieren de un nivel a otro sólo cuando se requiere. Además de reducir existencias, este enfoque permite a los trabajadores localizar con rapidez problemas de control de calidad. Como las partes se desplazan a través del sistema en pequeños lotes, los defectos se pueden identificar con rapidez. Por la naturaleza del sistema, cuando se detiene la producción en un lugar, automáticamente la producción en toda la línea se detiene de tal manera que la fuente del problema se pueda identificar y corregir antes de que se acumulen inventarios de artículos defectuosos.

Un aspecto importante de JIT es la reducción de los tamaños de lote de la producción. Con el fin de hacer más económicos los tamaños de lote, es necesario reducir los costos fijos del cambio. Ése es el objetivo del cambio de dados en un minuto (SMED). Si se divide la operación de cambio de dado o herramienta entre porciones que puedan realizarse fuera de la línea y aquellas que puedan llevarse a cabo en la línea, se pueden conseguir enormes reducciones en los tiempos de instalación.

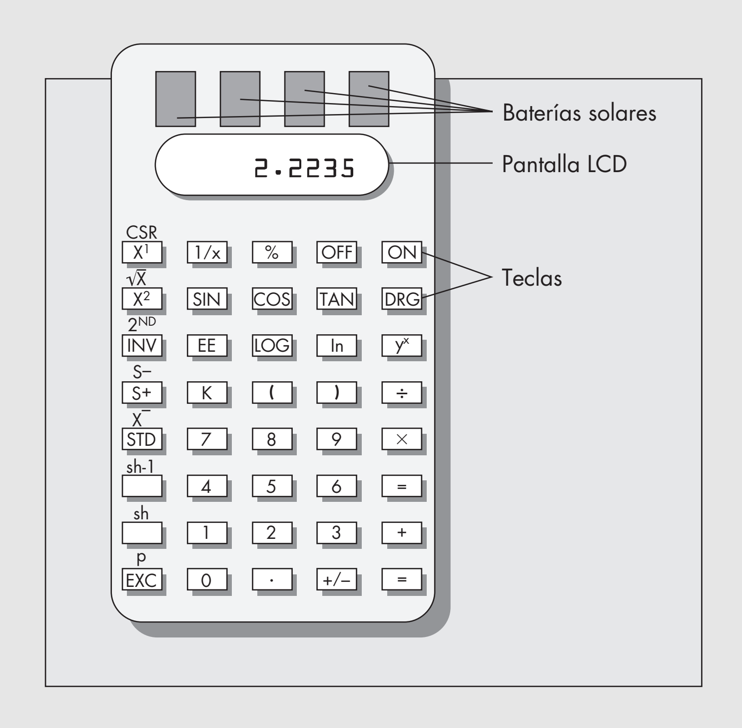
Kanban y JIT se encuentran estrechamente relacionados. Kanban es un sistema de información manual para implementar el JIT, el cual se basa en tarjetas o etiquetas para señalar la necesidad de más productos. Aunque no es de muy alta tecnología, el método ha funcionado bien en la práctica.

En comparación con MRP, JIT posee ventajas, así como desventajas. JIT reaccionará más rápidamente si se presenta un problema. Sin embargo, los sistemas JIT son muy lentos para reaccionar a los cambios en el esquema de demanda. Por otra parte, MRP construye pronósticos directamente en el cálculo de la explosión.

Problemas adicionales para el capítulo 7

41. CalcIt produce una línea de calculadoras de bolsillo económicas. El modelo IT53 es un modelo científico de energía solar con pantalla de cristal líquido (LCD). La calculadora se representa en la figura 7-11.

FIGURA 7-11
Calculadora científica CalcIt modelo IT53 (problema 41).



Cada calculadora requiere cuatro celdas solares, 40 teclas, una pantalla LCD y un procesador principal. Todas las partes se pidieron a un proveedor externo, pero el ensamble final lo lleva a cabo CalcIt. Los procesadores deben estar en bodega tres semanas antes de la fecha anticipada de terminación del lote de calculadoras para contar con suficiente tiempo para instalar el procesador en la caja, conectar cables y dejar que el pegamento seque. Las teclas deben estar en bodega con dos semanas de anticipación y se instalan a mano en las calculadoras. Las pantallas LCD y las celdas solares se solicitan al mismo proveedor y necesitan almacenarse una semana antes.

Sobre la base de los pedidos de la firma que CalcIt ha hecho, el programa maestro de producción para IT53 en un periodo de 10 semanas comenzando la semana 8 está dado por la siguiente tabla:

Semana	8	9	10	11	12	13	14	15	16	17
MPS	1 200	1 200	800	1 000	1 000	300	2 200	1 400	1 800	600

Determine el programa de requerimientos brutos para las celdas solares, las teclas, la pantalla LCD y los chips del procesador principal.

42. Considere el ejemplo de la compañía CalcIt en el problema 41. Suponga que las teclas de las calculadoras tienen un costo de $0.02 cada una y la compañía calcula un costo fijo de $12 por colocar y recibir pedidos de teclas a un proveedor externo. Suponga que los costos de almacenamiento se basan en una tasa de interés anual del 24 por ciento y que un año tiene 48 semanas. Aplicando el programa de requerimientos brutos para las teclas determinado en el problema 41, ¿qué política de pedidos recomienda la heurística Silver-Meal para las teclas? (Sugerencia: exprese h como costo por cada 10 000 unidades y divida cada demanda entre 10 000.)
43. Resuelva el problema 42 empleando el balanceo de periodos parciales y el costo unitario mínimo. Compare los costos de las soluciones resultantes con el costo de la solución obtenida utilizando la heurística Silver-Meal.
44. *Inventario de productos en proceso* (WIP) es un término que se refiere al inventario de componentes y ensambles en un proceso de manufactura. Suponiendo un programa lote

por lote en todos los niveles, el inventario WIP debe ser teóricamente igual a cero. ¿Piensa el lector que es probable que éste sea el caso de un entorno de manufactura real?

45. Vivian Lowe planea una fiesta sorpresa para el cumpleaños número 50 de su esposo. Ha decidido servir *sish kabob.* La receta requiere dos trozos de piña por cada camarón. Vivian planea racionar los *kabobs* de tal manera que cada uno tenga tres camarones. Ella calcula que de una sola piña saldrán unos 50 trozos, pero, por experiencia, sabe que aproximadamente una de cada 10 piñas salen malas y tiene que tirarlas. Invita a 200 personas y espera que cerca de la mitad se presente. Cada persona generalmente se come 2 *kabobs.*

   a) ¿Cuántas piñas debe comprar?

   b) Supongamos que el número de invitados es una variable aleatoria con una distribución normal con una media de 100 y una varianza de 1,680. Si desea elaborar suficientes *kabobs* para alimentar a todos los invitados con una probabilidad del 95 por ciento, ¿cuántas piñas debería comprar?

46. En este capítulo supusimos que la "ventana de tiempo" era de una semana. Esto implica que los pronósticos se reevalúen y que el sistema MRP vuelva a correr semanalmente.

   a) Comente las posibles ventajas de emplear una ventana de tiempo más corta, por ejemplo, un día.

   b) Comente las posibles ventajas de emplear una ventana de tiempo más larga, como dos semanas o un mes.



47. Elabore una hoja de cálculo que reproduzca los cálculos del ejemplo 7.1. Como en el ejemplo, la hoja de cálculo debe incluir la demanda neta predicha para las trompetas. Las columnas deben corresponder a semanas y deben etiquetarse del 1 al 18. Debajo de la demanda neta predicha para las trompetas deben colocarse los cálculos para el ensamble de la caja de válvulas y debajo los cálculos para las válvulas. En el caso de cada componente, incluya los renglones para la siguiente información: 1) requerimientos brutos; 2) entradas programadas; 3) existencias disponibles; 4) requerimientos netos trasladados; 5) requisiciones planeadas lote por lote. La hoja de cálculo debe actualizar automáticamente todos los cálculos si la demanda neta predicha para las trompetas cambia.



48. Dos productos terminales, EP1 y EP2, se producen en la planta de Raleigh, Carolina del Norte, de una gran fábrica de productos para muebles localizada en el sureste. Los diagramas de estructura de los productos aparecen en la figura 7-12.

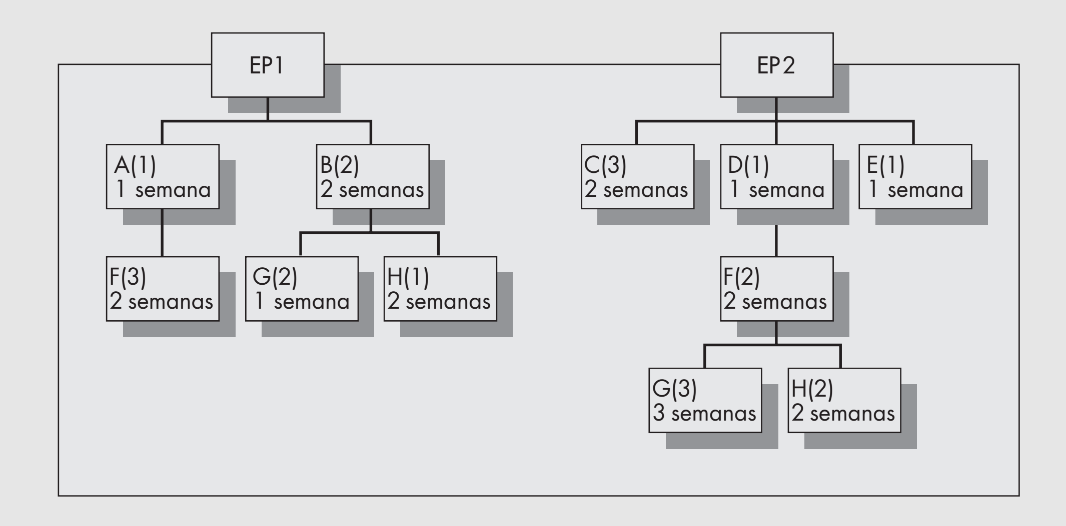


FIGURA 7-12
Diagramas de estructura del producto (problema 48).

Suponga que los programas maestros de producción para estos dos productos son:

Semana	18	19	20	21	22	23	24
EP1	120	112	76	22	56	90	210
EP2	62	68	90	77	26	30	54

Suponiendo una producción lote por lote, determine las requisiciones planeadas para los componentes F, G y H.

49. Un componente empleado en cierto equipo de fabricación se solicita a un proveedor externo. Como el componente se utiliza en una variedad de productos finales, la demanda es alta. La demanda aproximada (en miles) las siguientes 10 semanas es:

Semana	1	2	3	4	5	6	7	8	9	10
Demanda	22	34	32	12	8	44	54	16	76	30

El costo de los componentes es de 65 centavos y la tasa de interés empleada para calcular el costo de almacenamiento es de 0.5 por ciento semanal. El costo fijo de pedido se calcula en $200. (Sugerencia: exprésese h como costo de almacenamiento por cada mil unidades.)

a) ¿Qué política de pedidos recomienda la heurística de Silver-Meal?
b) ¿Qué política de pedidos recomienda la heurística de balanceo de periodos parciales?
c) ¿Qué política recomienda la heurística de costo mínimo unitario?
d) ¿Con cuál método se obtuvo la política de costo más bajo en este problema?

50. Un método heurístico popular para determinar tamaños de lote recibe el nombre de cantidad periódica de pedidos (POQ, *period order quantity*). El método requiere determinar el promedio de periodos generados por la CEP y elegir el tamaño de lote para igualar este periodo fijo de suministro. Sea λ la demanda total durante el horizonte de planeación de N periodos $[\lambda = (\Sigma\, r_i/n)]$ y suponga que la CEP se calcula según se describe en la sección 7.2. Entonces $P = \text{EOQ}/\lambda$, redondeado al entero más próximo. Para el ejemplo de la sección 7.2, $P = 139/43.9 = 3.17$, que se redondea a 3. El POQ requeriría igualar el tamaño del lote a tres periodos de demanda. Para el ejemplo de la sección 7.2, las requisiciones planeadas resultantes serían de 116, 0, 0, 150, 0, 0, 135, 0, 0, 38.

a) Compare el costo de la política obtenido por medio de este método en el caso del ejemplo de la sección 7.2 con el que se obtuvo con el método CEP.
b) ¿Cuáles son las ventajas de este enfoque sobre CEP?
c) ¿Piensa usted que este método resultará, en general, más eficaz en lo que se refiere a costos que los métodos heurísticos analizados en la sección 7.2?
d) Resuelva el problema 17 empleando este método y compare el costo total de instalación y almacenamiento con el costo obtenido por otros métodos.
e) Resuelva el problema 49 empleando este método y compare el costo total de instalación y almacenamiento con el costo obtenido por otros métodos.

51. La tienda del una gran universidad de la región central de Estados Unidos vende cuadernos a los estudiantes, profesores y administrativos. Venden más cuadernos al aproximarse el tiempo de exámenes. Durante un trimestre típico de 10 semanas, el esquema de ventas es de 2 280, 1 120, 360, 3 340, 1 230, 860, 675, 1 350, 4 600, 1 210. Los cuadernos cuestan a la tienda $1.20 cada uno, y los costos de almacenamiento tienen una tasa de interés anual del 30 por ciento. El costo del tiempo por empleado, trámites y manejo suman $30 por pedido. Supóngase que el año tiene 50 semanas.

a) ¿Cuál es la política óptima de pedidos durante el trimestre de 10 semanas con base en la heurística de Silver-Meal? De acuerdo con esta política, ¿cuáles son los costos de almacenamiento y pedidos en los que se incurre durante el periodo de 10 semanas?

b) El gerente de la tienda decide que resultaría más económico que la demanda fuera la misma cada semana. Con el fin de unificar la demanda limita las ventas semanales (para sorpresa de sus clientes). Así, supone que la demanda total para el trimestre de 10 semanas sea la misma, aunque las ventas son constantes semana a semana. Determine la política óptima de pedidos en este caso y compare el costo total de almacenamiento y preparación durante las 10 semanas con la respuesta obtenida en el inciso *a*). (Se puede suponer un tiempo continuo para realizar los cálculos, así que el tamaño de lote óptimo es el CEP.)

c) Con base en los resultados de los incisos *a*) y *b*), ¿piensa usted que en general resulta más económico tener un esquema de demanda estable o uno con picos?



52. Desarrolle una hoja de cálculo que permita determinar soluciones Silver-Meal para problemas generales de determinación de tamaños de lote. Guarde los parámetros de costo de almacenamiento e instalación en celdas separadas con el fin de que se puedan introducirse y cambiarse valores con facilidad. Introduzca en la columna 2 de la hoja de cálculo 30 periodos de demanda. Enliste los números de periodo 1, 2, . . . , 30 en la columna 1. Ejecute la lógica que proporciona $C(j)$ en la columna 3.

La hoja se utilizaría de la siguiente manera: introduzca los requerimientos periodo por periodo hasta que se observe un incremento en la columna 3. Esto identifica el primer horizonte de pronósticos. Después remplace las entradas de la columna 2 por ceros e introduzca los requerimientos comenzando en el periodo actual. Continúe hasta que el siguiente horizonte de pronósticos se identifique. Repita el procedimiento hasta el final del horizonte de planeación. Utilice este método para determinar la solución de Silver-Meal en el caso de los siguientes problemas de planeación de la producción:

a) Resuelva el problema 14 con este procedimiento.

b) Se predice que las demandas semanales de remaches de 2 pulgadas en una división de una compañía de aviones son las siguientes (en gruesas):

Semana	1	2	3	4	5	6	7	8	9	10	11	12
Demanda	240	280	370	880	950	120	135	450	875	500	400	200
Semana	13	14	15	16	17	18	19	20	21	22	23	24
Demanda	600	650	1,250	250	800	700	750	200	100	900	400	700

Los costos de preparación por ordenar remaches se calcula que son de $200, y los costos de almacenamiento ascienden a 10 centavos por gruesa cada semana. Determine el tamaño de lote que se obtiene con el método de Silver-Meal.



53. Con las instrucciones del problema 52, construya una hoja de cálculo para determinar la regla de dimensionamiento de tamaños de lote con costo unitario mínimo.

a) Resuelva los incisos *a*) y *b*) del problema 52.

b) ¿Cuál método, el de costo unitario mínimo o el de Silver-Meal, proporciona una solución más eficaz en cuanto a costos?

Apéndice 7-A

Determinación óptima de tamaños de lote para demandas que varían con el tiempo

Las técnicas de la sección 7.3 son fáciles de aplicar y proporcionan tamaños de lote con costos que por lo general se encuentran cerca de ser auténticamente óptimos. Este apéndice explica la forma de calcular los tamaños de lote realmente óptimos. En este contexto, *óptimo* implica

la política por medio de la cual se minimiza el costo total de preparación y almacenamiento en un horizonte de planeación. Este apéndice muestra la forma en que las políticas óptimas pueden determinarse convirtiéndolo en un problema de encontrar el camino más corto.

Suponga lo siguiente:

1. Las demandas pronosticadas durante los siguientes n periodos se conocen y están dadas por el vector $\mathbf{r} = (r_1, \ldots, r_n)$.
2. Los costos de almacenamiento son de $\$h$ por unidad por periodo y $\$K$ por preparación. Supondremos que el costo de instalación se carga al final del inventario cada periodo.

Con el fin de obtener una idea de las posibles dificultades del problema, consideremos el siguiente problema sencillo.

Ejemplo *7A.1*

La demanda pronosticada para un aparato electrónico hecho en un taller de fabricación de semiconductores de alta tecnología local durante las siguientes cuatro semanas es de 52, 87, 23, 56. Sólo hay una preparación por semana para la producción de estos aparatos, y no hay pedidos pendientes de demanda en exceso. Supongamos que el taller posee la capacidad de producir cualquier cantidad de aparatos en una semana.

Consideremos el número total de políticas de producción para el aparato durante el periodo de cuatro semanas. Sean $y_1, \ldots, y_4$ las cantidades ordenadas en cada una de las cuatro semanas. Evidentemente, $y_1 \geq 52$, para garantizar el hecho de que no hagan falta existencias el periodo 1. Si suponemos que el inventario final en el periodo 4 es igual a cero (sería fácil demostrar que es óptimo), entonces $y_1 \leq 218$, la suma de todas las demandas. De ahí que y_1 pueda adoptar cualquiera de los 167 posibles valores. Consideremos y_2. El número de valores factibles de y_2 depende del valor de y_1. Se permite que ocurra un agotamiento de existencias en el periodo 2, $y_1 + y_2 \geq 52 + 87 = 139$. Si

$$y_1 \leq 139, \qquad \text{entonces } 139 - y_1 \leq y_2 \leq 218 - y_1,$$

y si

$$y_1 > 139, \qquad \text{entonces } 0 \leq y_2 \leq 218 - y_1.$$

Con un poco de esfuerzo es posible demostrar que esto da como resultado un total de 10,200 diferentes valores exclusivamente de la pareja (y_1, y_2). Por consiguiente, resulta evidente que incluso en problemas de dimensiones moderadas, el número de soluciones factibles es enorme.

La búsqueda de todas las políticas factibles resulta ilógica. Sin embargo, un importante descubrimiento de Wagner y Within reduce considerablemente la cantidad de políticas que deben considerarse como candidatas para la optimización.

El algoritmo de Wagner-Within se basa en la siguiente observación:

Resultado. Una política óptima posee la propiedad de que cada valor de **y** es exactamente la suma de un conjunto de demandas futuras. (Llamaremos a esto una *política de requerimientos exactos.*) Es decir que

$$y_1 = r_1, \qquad \text{o } y_1 = r_1 + r_2, \ldots, \qquad \text{or } y_1 = r_1 + r_2 + \cdots + r_n.$$

$$y_2 = 0, \qquad \text{o } y_2 = r_2, \qquad \text{o } y_2 = r_2 + r_3, \ldots, \qquad \text{o}$$

$$y_2 = r_2 + r_3 + \cdots + r_n$$

$$\vdots$$

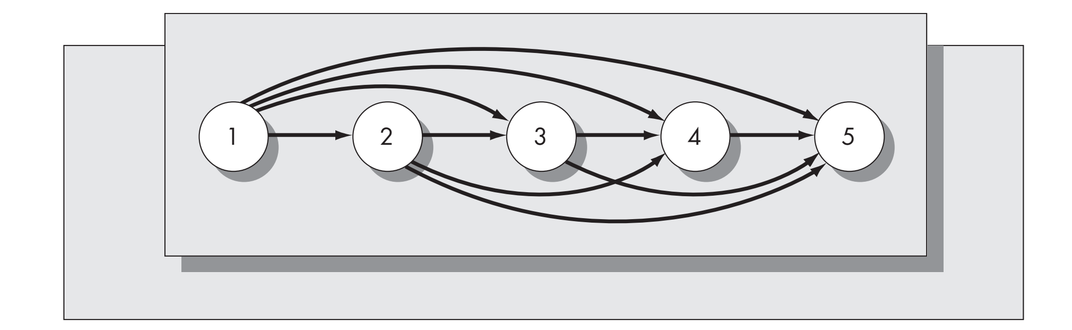
$$y_n = 0 \qquad \text{o } y_n = r_n.$$

Una política de requerimientos exactos se especifica completamente designando los periodos en los que deben realizarse los pedidos.

Ejemplo *7.A.1 (continuación)*

Continuemos con el problema de programación de cuatro periodos. Como y_1 debe satisfacer requerimientos exactos, vemos que únicamente puede tomar valores de 52, 139, 162 o 218; es decir, sólo cuatro diferentes valores. Si no se toma en cuenta el valor de y_1, y_2 puede adoptar los valores 0, 87, 110, 166. Resulta fácil ver que cada política de requerimientos exactos se determina completamente especificando en qué periodos deben llevarse a cabo los pedidos. Es decir que cada política tiene la forma $(i_1, \ldots, i_n)$, en la que los valores de i_j son 0 o 1.

FIGURA 7A.1
Representación en red para la programación de lotes (ejemplo 7A.1).



Si $i_j = 1$, entonces la producción se realiza en el periodo j. Observe que $i_1 = 1$ porque debemos producir en el periodo 1 para evitar quedarnos sin existencias, mientras que $i_2, \ldots, i_n$ serán cada uno 0 o 1. Por ejemplo, la política (1, 0, 1, 0) significa que la producción ocurre los periodos 1 y 3 solamente. Se deduce que $\mathbf{y} = (139, 0, 79, 0)$. En este ejemplo, hay exactamente $2^3 = 8$ diferentes políticas de requerimientos exactos.

Una manera conveniente de ver el problema es como si se tratara de una red en un solo sentido con un número de nodos exactamente igual al número de periodos más uno. Cada trayectoria a través de la red corresponde a una política de requerimientos exactos. La red para el problema de cuatro periodos aparece en la figura 7A-1.

Para cualquier pareja (i, j) con $i < j$, si el arco (i, j) se encuentra en la trayectoria, esto significa que los pedidos tienen lugar en el periodo i y que el tamaño del pedido es igual a la suma de los requerimientos en los periodos $i, i + 1, \ldots, j - 1$. El periodo j es el siguiente periodo de pedido. Obsérvese que todas las trayectorias terminan en el periodo $n + 1$. La política de pedidos en los periodos 1 y 3 solamente correspondería a la ruta 1–3–5. La ruta 1–2–4–5 implica que el pedido tendrá lugar los periodos 1, 2 y 4.

El siguiente paso es asignar un valor a cada arco de la red. El valor o *longitud* del arco (i, j), denominado c_{ij}, se define como el costo de almacenamiento y preparación del pedido en el periodo i para cubrir las demandas durante el periodo $j - 1$. Por ejemplo, c_{15} = al costo de ordenar en el periodo 1 para satisfacer las demandas de los periodos 1 a 4.

Por último, quisiéramos determinar el programa de producción de costo mínimo, o la ruta más corta a través de la red. Como veremos, la *programación dinámica* es un método para resolver este problema. Sin embargo, en un problema pequeño, es posible determinar la política óptima sencillamente enumerando los caminos a través de la red y eligiendo uno con costo mínimo.

Ejemplo *7.A.2*

Resolveremos el ejemplo 7A.1 aplicando la enumeración de trayectorias. Recordemos que $\mathbf{r} = (52, 87, 23, 56)$. Además, supongamos que existe un costo de almacenamiento de $h = \$1$ por unidad por periodo y un costo de $K = \$75$ por preparación.

El primer paso consiste en calcular c_{ij} para $1 \le i \le 4$ y $i + 1 \le j \le 5$.

$c_{12} = 75$ (sólo costo de preparación).

$c_{13} = 75 + 87 = 162.$

$c_{14} = 75 + (23 \times 2) + 87 = 208.$

$c_{15} = 75 + (56 \times 3) + (23 \times 2) + 87 = 376.$

$c_{23} = 75.$

$c_{24} = 75 + 23 = 98.$

$c_{25} = 75 + 23 + (56 \times 2) = 210.$

$c_{34} = 75.$

$c_{35} = 75 + 56 = 131.$

$c_{45} = 75.$

Resumiendo estos costos en forma de matriz se obtiene lo siguiente:

i \ *j*	1	2	3	4	5
1		75	162	208	376
2			75	98	210
3				75	131
4					75

Como sólo hay ocho políticas de requerimientos exactos, podemos resolver este problema enumerando las políticas y comparando costos.

Trayectoria	Costo
1–2–3–4–5	$300
1–2–4–5	248
1–2–5	285
1–2–3–5	281
1–3–4–5	312
1–3–5	293
1–4–5	283
1–5	376

Se deduce que la trayectoria óptima es 1–2–4–5 a un costo de $248. Esto corresponde a pedir sólo en los periodos 1, 2 y 4. La política óptima de pedidos es $y_1 = 52$, $y_2 = 110$, $y_3 = 0$, $y_4 = 56$.

*SOLUCIÓN CON PROGRAMACIÓN DINÁMICA

El número total de políticas de requerimientos exactos para un problema de n periodos es de 2^{n-1}. Conforme n aumenta, la enumeración total no es eficiente. La programación dinámica es una técnica de solución recurrente que reduce significativamente el número de cálculos que se requieren, aunque también puede resultar complicada.

La programación dinámica se basa en el *principio de optimalidad.* Una versión de este principio es la que menciona que si un problema consta exactamente de n etapas y quedan $r < n$ etapas, la política óptima para las etapas restantes es independiente de la política adoptada en las etapas anteriores. Como la programación dinámica no se emplea en cualquier parte en este texto, no la explicaremos con detalle. El lector interesado puede consultar a Hillier y Lieberman (1990), que incluyen un breve panorama, o a Nemhauser (1966), que contiene un estudio más profundo con un nivel matemático más congruente con el nuestro.

Definamos f_k como el costo mínimo comenzando en el nodo k, suponiendo que se coloca un pedido en el periodo k. El principio de optimalidad para este problema da como resultado el siguiente sistema de ecuaciones:

$$f_k = \min_{j>k}(c_{kj} + f_j) \qquad \text{para } k = 1, \ldots, n.$$

La condición inicial es $f_{n+1} = 0$.

Ejemplo *7A.3* Resolveremos el ejemplo 7A.1 con programación dinámica para ilustrar la técnica. Comencemos con la condición inicial y regresemos del periodo $n + 1$ al periodo 1.[4] En cada periodo se determina el valor de j que permite obtener el mínimo.

[4] En realidad, el algoritmo original de Wagner-Within (1958) se basa en una formulación de programación dinámica progresiva. Aunque la formulación progresiva posee algunas ventajas para el análisis del horizonte de planeación, creemos que la recursividad regresiva es más natural e intuitiva.

$$f_5 = 0.$$

$$f_4 = \min_{j>4}(c_{4j} + f_j)$$

$$= 75 \qquad \text{en } j = 5 \text{ (el único valor posible de } j\text{).}$$

$$f_3 = \min_{j>3} (c_{3j} + f_j) = \text{mín} \begin{Bmatrix} c_{34} + f_4 \\ c_{35} + f_5 \end{Bmatrix} = \text{mín} \begin{Bmatrix} 75 + 75 \\ 131 + 0 \end{Bmatrix} = \text{mín} \begin{Bmatrix} 150 \\ 131 \end{Bmatrix}$$

$$= 131 \qquad \text{en } j = 5.$$

$$f_2 = \min_{j>2} (c_{2j} + f_j) = \text{mín} \begin{Bmatrix} c_{23} + f_3 \\ c_{24} + f_4 \\ c_{25} + f_5 \end{Bmatrix} = \text{mín} \begin{Bmatrix} 75 + 131 \\ 98 + 75 \\ 210 + 0 \end{Bmatrix} = \text{mín} \begin{Bmatrix} 206 \\ 173 \\ 210 \end{Bmatrix}$$

$$= 173 \qquad \text{en } j = 4.$$

Por último,

$$f_1 = \min_{j>1} (c_{1j} + f_j) = \text{mín} \begin{Bmatrix} c_{12} + f_2 \\ c_{13} + f_3 \\ c_{14} + f_4 \\ c_{15} + f_5 \end{Bmatrix} = \text{mín} \begin{Bmatrix} 75 + 173 \\ 162 + 131 \\ 208 + 75 \\ 376 + 0 \end{Bmatrix} = \text{mín} \begin{Bmatrix} 248 \\ 293 \\ 283 \\ 376 \end{Bmatrix}$$

$$= 248 \qquad \text{en } j = 2.$$

Para determinar la política de pedidos óptima, volvemos a proyectar la solución hacia atrás desde el inicio. En el periodo 1, el valor óptimo de j es $j = 2$. Esto significa que el nivel de producción en el periodo 1 es igual a la demanda en el periodo 1, así que $y_1 = r_1 = 52$. El siguiente periodo de pedido es el periodo 2. El valor óptimo de j en el periodo 2 es $j = 4$, lo cual implica que la producción en el periodo 2 es igual a la suma de las demandas en los periodos 2 y 3, o $y_2 = r_2 + r_3 = 110$. El siguiente periodo de pedido es el periodo 4. El valor óptimo de j en el periodo 4 es $j = 5$. Esto da como resultado $y_4 = r_4 = 56$. Por lo tanto, la política óptima de pedidos es $\mathbf{y} = (52, 110, 0, 56)$.

Apéndice 7.B

Glosario de notación para el capítulo 7

$C(T)$ = Costo promedio de almacenamiento y preparación por periodo (para la heurística de Silver-Meal) o por unidad (heurística LUC) si el pedido actual abarca T periodos.

c_i = Capacidad de producción en el periodo i.

c_{ij} = Costo asociado con el arco (i, j) en la representación en red del problema de programación del lote empleado para el algoritmo de Wagner-Within.

f_j = Costo mínimo del periodo i al final del horizonte (consulte el algoritmo de programación dinámica para Wagner-Within).

h = Costo de almacenamiento por unidad por periodo de tiempo.

K = Costo de preparación para dar inicio a un pedido.

r_i = Requerimiento para el periodo i.

y_i = Tamaño de lote de la producción el periodo i.

Bibliografía

Afentakis, P. B., Gavish y U. Karmarkar. "Computationally Efficient Optimal Solutions to the Lot-Sizing Problem in Multistage Assembly System". *Management Science* 30, 1984, pp. 222-239.

Anderson, J. C., R. G. Schroeder, S. E. Tupy y E. M. White. "Material Requirements Planning Systems: The State of the Art". *Production and Inventory Management* 23, 1982, pp. 51-56.

Blood B. E. "Read My LipsæNo More Late Deliveries". *Hospital Material Management Quarterly,* 15, núm. 4, 1994, pp. 53-55.

Carlson, R. C., S. L. Beckman y D. H. Kropp. "The Effectiveness of Extending the Horizon in Rolling Production Scheduling". *Decision Sciences* 13, 1982, pp. 129-146.

Carlson, R. C., J. V. Jucker y D. H. Kropp. "Less Nervous MRP Systems: A Dynamic Economic Lot-Sizing Approach". *Management Science* 25, 1979, pp. 754-761.

Carlson, R. C., J. V. Jucker y D. H. Kropp. "Heuristic Lot Sizing Approaches for Dealing with MRP System Nervousness". *Decision Sciences* 14, 1983, pp. 159-169.

Chhikara, J. y E. N. Weiss. "JIT Savings—Myth or Reality?" *Business Horizons* 38, mayo-junio, 1995, pp. 73-78.

DeMatteis, J. J. "An Economic Lot Sizing Technique: The Part-Period Algorithm". *IBM Systems Journal* 7, 1968, pp. 30-38.

Dixon, P. S. y E. A. Silver. "A Heuristic Solution Procedure for the Multi-Item, Single-Level, Limited Capacity Lot Sizing Problem". *Journal of Operations Management* 2, 1981, pp. 23-39.

Gorham, T. "Dynamic Order Quantities". *Production and Inventory Management* 9, 1968, pp. 75-81.

Hillier, F. S. y G. J. Lieberman. *Introduction to Operations Research.* 5a. ed., Nueva York: McGraw-Hill, 1990.

Hopp, W. y M. Spearman. *Factory Physics.* Burr Ridge, IL: Richard D. Irwin, 1996.

Jacobson, G. y J. Hillkirk. *Xerox, American Samurai.* Nueva York: MacMillan, 1986.

Karmarkar, U. "Getting Control of Just-in-Time". *Harvard Business Review* 67, septiembre-octubre, 1989, pp. 122-131.

Karni, R. "Maximum Part Period Gain (MPG)—A Lot Sizing Procedure for Unconstrained and Constrained Requirements Planning Systems". *Production and Inventory Management* 22, 1981, pp. 91-98.

Krajewski, L. J., B. E. King, L. P. Ritzman y D. S. Wong. "Kanban, MRP, and Shaping the Manufacturing Environment". *Management Science* 33, 1987, pp. 39-57.

LaPlante, A. "New Software = Faster Factories". *Forbes ASAP Supplement*, octubre 10, 1994, pp. 36-41.

Lieber, R. B. "Here Comes SAP". *Fortune* 132, núm. 7, octubre 10, 1995, pp. 122-124.

Love, Stephen. *Inventory Control*, Nueva York, McGraw-Hill, 1979.

McClaren, B. J. "A Study of Multiple Level Lot Sizing Techniques for Material Requirements Planning Systems". Tesis doctoral inédita, Universidad de Purdue, 1976.

McLeavey, D. W. y S. L. Narasimhan. *Production Planning and Inventory Control.* Boston: Allyn & Bacon, 1985.

Meal, H. "Putting Production Decisions Where They Belong". *Harvard Business Review* 62, 1984, pp. 102-111.

Monden, Y. "What Makes the Toyota Production System Really Tick?" *Industrial Engineering* 13, núm. 1, 1981a, pp. 36-46.

Monden, Y. "Adaptable Kanban Systems Help Toyota Mantain Just-in-Time Production". *Industrial Engineering* 13, núm. 5, 1981b, pp. 28-46.

Nahmias, S. y K. Moinzadeh. "Lot Sizing with Randomly Graded Yields". *Operations Research* 46, núm. 6, 1997, pp. 974-986.

Nemhauser, G. L. *Introduction to Dynamic Programming,* Nueva York, John Wiley & Sons, 1996.

New, C. *Requirements Planning,* Essex, Inglaterra: Gower Press, 1974.

Orlicky, J. *Materials Requirements Planning*, Nueva York: McGraw-Hill, 1975.

Plossl, G. y O. Wight. *Materials Requirements Planning by Computer*, Washington, DC: American Production and Inventory Control Society, 1971.

Sheldon, D. "MRP II Implementation: A Case Study". *Hospital Material Management Quarterly* 15, núm. 4, 1994, pp. 48-52.

Shingo, S. *Study of "Toyota" Production System from Industrial Engineering Viewpoint.* Tokio: Japan Management Association, 1981.

Silver, E. A. y H. C. Meal. "A Heuristic for Selecting Lot Size Quantities for the Case of a Deterministic Time-Varying Demand Rate and Discrete Opportunities for Replenishment". *Production and Inventory Management* 14, 1973, pp. 64-74.

Silver, E. A. y R. Peterson. *Decision Systems for Inventory Management and Production Planning.* 2a. ed. Nueva York: John Wiley & Sons, 1985.

Sloan, A. *My Years with General Motors,* Garden City, NY: Doubleday, 1964.

Steele, D. C. "The Nervous MRP System: How to Do Battle". *Production and Inventory Management* 16, 1973, pp. 83-89.

Steinberg, E. y H. A. Napier. "Optimal Multi-Level Lot Sizing for Requirements Planning Systems". *Management Science* 26, 1980, pp. 1258-1271.

Vollman, T. E., W. L. Berry y D. C. Whybark. *Manufacturing and Control Systems*. 3a. ed. Nueva York, McGraw-Hill/Irwin, 1992.

Wagner, H. M. y T. M. Within. "Dynamic Version of the Economic Lot Size Model". *Management Science* 5, 1958, pp. 89-96.

Willis, R. "Harley Davidson Comes Roaring Back". *Management Review* 75, marzo de 1986, pp. 20-27.

Womack, J. P., D. T. Jones y D. Roos. *The Machine That Changed the World: The Story of Lean Production.* Nueva York: Harper Perennial, 1990.

Zipkin, P. "Does Manufacturing Need a JIT Revolution?" *Harvard Business Review* 69, enero-febrero, 1991, pp. 40-50.

Capítulo **Ocho**

Programación de operaciones

Panorama general del capítulo

Objetivo

Entender los métodos esenciales y los resultados de la programación secuencial en el contexto de trabajo en un taller.

Puntos clave

1. *El problema de la programación en un taller.* Un taller esta conformado por máquinas de trabajo y personas encargadas de su operación. Los trabajos pueden llegar todos a la vez o de manera aleatoria durante el día. Por ejemplo, considere un lugar donde se reparan automóviles. Para un día dado, no puede predecirse exactamente qué tipos de reparaciones van a llegar al taller. Cada trabajo requiere diferentes equipos y posiblemente distinto personal. A un mecánico con experiencia se le puede asignar un trabajo complicado, como cambiar una transmisión, mientras que a uno novato se le asignaría el mantenimiento de rutina. Suponga que los clientes traen sus automóviles a primera hora de la mañana. El capataz del taller necesita determinar la secuencia de programación de los trabajos para lograr un aprovechamiento más eficiente de los recursos disponibles (tanto humanos como mecánicos).

   Las características relevantes del problema de secuenciación incluyen

   - El patrón de llegadas.
   - El número y la variedad de las máquinas.
   - El número y el tipo de los trabajadores.
   - Los patrones de flujo de trabajo en el taller.
   - Los objetivos de evaluar las reglas alternativas de secuenciación.

2. *Las reglas de secuenciación.* En esta sección se consideran las siguientes reglas:

   - *Primeras llegadas, primeras salidas (FCFS: first-come, first send).* Programe los trabajos en el orden de llegada al taller.
   - *Tiempo de procesamiento más corto primero (SPT: shortest processing time).* Programe primero el trabajo con el tiempo de procesamiento más corto.
   - *Primera fecha de entrega (EDD: earliest due date).* Programe primero los trabajos que tengan la fecha de entrega más cercana.
   - *Programación basada en la razón crítica (CR: critical ratio).* La razón crítica es (fecha de entrega – fecha actual)/tiempo de procesamiento. Programe en seguida el trabajo con el valor más pequeño de CR.

3. *Resultados de la secuenciación.* Un criterio común para evaluar la efectividad de las reglas de secuenciación es el tiempo de flujo medio. El tiempo de flujo de cualquier trabajo es el lapso que transcurre desde el momento en que el trabajo llega al taller, hasta que el trabajo se completa. El tiempo de flujo medio es simplemente el promedio de todos los tiempos de flujo de todos los trabajos. El principal resultado de esta sección es que la programación SPT minimiza el tiempo de flujo medio. Otra interesante consecuencia sería que, si el objetivo consiste en minimizar el retraso máximo, entonces los trabajos deben programarse mediante EDD. Esta sección también estudia varios algoritmos de programación. El algoritmo de Moore aminora el número de trabajos retrasados, y el algoritmo de Lawler se usa cuando están presentes restricciones de precedencia (es decir, que los trabajos deben hacerse según cierto orden).

   Todos los resultados anteriores se aplican a una sola máquina o a una sola instalación. Al programar trabajos para máquinas múltiples, el problema es mucho más complejo. En tales instancias, hay algunos resultados conocidos. Considere el ejemplo de *n* trabajos que deben programarse para dos máquinas. El principal resultado descubierto en este caso es que la solución óptima consiste en secuenciar los trabajos en el mismo orden para ambas máquinas (lo cual se conoce como programa de permutación). Esto significa que existen *n*! soluciones factibles posibles. Por supuesto que éste puede ser un número muy grande. Sin embargo, un procedimiento descubierto por Johnson (1954) calcula eficientemente la secuencia óptima de *n* trabajos para dos máquinas. Puede aplicarse el mismo algoritmo a tres máquinas bajo circunstancias muy especiales. El problema de programar dos trabajos para *m* máquinas puede resolverse eficientemente mediante un procedimiento gráfico.

4. *La programación secuencial en un ambiente estocástico.* Los problemas que se mencionan anteriormente suponen que toda la información se conoce con certeza. Los problemas reales son más complejos pues generalmente existe algún tipo de incertidumbre, como podrían ser las duraciones de los trabajos. En ese caso, se supone que éstas, digamos $t_1, t_2, \ldots, t_n$, son variables aleatorias independientes con una función de distribución α conocida. La secuencia óptima para una sola máquina en este caso se parece mucho a la programación de los trabajos en el orden SPT con base en los tiempos de procesamiento esperados.

   Cuando se programan trabajos con tiempos de procesamiento inciertos en máquinas múltiples, debe suponerse que las duraciones de los trabajos se adhieren a una distribución exponencial. Ésta es la única que posee la propiedad sin memoria, crucial en el análisis. Si el objetivo es minimizar la terminación esperada (es decir, el tiempo total para acabar todos los trabajos), resulta que la primera regla del tiempo de procesamiento máximo esperado (LEPT: *longest expected processing time*) es óptima.

   Otra fuente de incertidumbre en un taller es el orden en que los trabajos son recibidos. En el ejemplo del taller automotriz, supusimos que los trabajos llegan todos a la vez al inicio del día. Sin embargo, en una fábrica es posible que lo hagan de manera aleatoria durante el día. En este caso, la teoría de colas puede arrojar alguna luz sobre cuánto tiempo transcurre desde el momento en que un trabajo llega y hasta su terminación. Esta sección esboza varios resultados bajo las hipótesis de las secuencias de FCFS, LCFS, y SPT.

5. *Balanceo de línea.* Otro problema que surge en un contexto fabril es el balanceo de una línea de ensamble. Aunque éste no es un problema de programación de secuencias que se presente en un taller, ciertamente es un problema de programación que surge dentro de la planta. Suponga que tenemos un artículo que fluye por una línea de ensamblado y que deben terminarse un total de n tareas sobre el artículo. El problema consiste en determinar dónde deben colocarse qué tareas en la línea. Es común que una línea de ensamblado se divida en estaciones y que algún subconjunto de tareas se asigne a cada una. El objetivo es balancear el tiempo requerido en cada estación tomando en cuenta

las relaciones de precedencia que existen entre las tareas individuales. Consideraremos un método heurístico, que proporciona resultados razonables en la mayoría de las circunstancias.

Después de este capítulo se encuentra un suplemento sobre la teoría de colas, que suministra un breve resumen de los resultados más importantes en esta área.

La programación es un aspecto importante del control de operaciones, tanto en la manufactura como en las industrias de servicios. Haciendo un gran énfasis en el tiempo de comercialización y en el tiempo para incrementar la producción, así como en la satisfacción del cliente, la programación eficiente va a adquirir auge en la función de las operaciones en los años venideros.

Gran parte de lo que se ha estudiado hasta ahora en este libro, en cierto modo puede considerarse como un subconjunto de la programación de la producción. La planeación agregada, que se estudia en el capítulo 3, está dirigida a la macro programación de los niveles de la fuerza de trabajo y de los niveles generales de producción de la compañía. El control detallado de los inventarios, que se estudia en los capítulos 4 y 5, trata sobre los métodos para programar la producción al nivel del elemento individual; el capítulo 6 ahonda en la programación de los vehículos; y el MRP, que se estudia en el capítulo 7, suministra calendarios de producción para artículos terminados y para subensamblados en la estructura del producto.

Una compañía enfrenta muchos tipos diferentes de problemas de programación. Una lista parcial incluye:

1. *Programación del taller*. Más conocida en la práctica como control de piso, es el conjunto de actividades que transforma los insumos (un conjunto de requerimientos) en salidas (productos que satisfacen a estos requerimientos). Gran parte de este capítulo está dedicado a la secuencia de aspectos del piso, y se va a hablar más acerca de este problema en la sección 8.1.
2. *Programación de personal*. Constituye un problema importante tanto para las industrias manufactureras como para las de servicio. Aun cuando la programación de los turnos de la fábrica puede considerarse como una de las funciones del control de piso, la programación de personal es un problema mucho mayor. Por ejemplo, el caso de los profesionales de la salud en los hospitales y en otras instalaciones de salud. Determinar cómo satisfacer la demanda pico ya sea con turnos de horas extras, turnos nocturnos, o la subcontratación es otro ejemplo de un problema de programación de personal.
3. *Programación de instalaciones*. Este problema es particularmente importante cuando las instalaciones se convierten en un recurso de cuello de botella. Muestra de ello sería la programación de las salas de operaciones de los hospitales. A medida que aumenta la necesidad del servicio de salud, algunos hospitales y organizaciones encuentran que la capacidad de sus instalaciones es rebasada. Un problema similar se presenta en las escuelas superiores y las universidades donde las inscripciones han crecido sin un aumento proporcional en el tamaño de la planta física.
4. *Programación de vehículos*. Las compañías manufactureras deben distribuir sus productos de una manera oportuna y eficiente en cuanto a costos. Algunas operaciones de servicio, tales como los sistemas telefónicos para solicitar transporte, incluyen recoger y entregar bienes y/o personas. Establecer las rutas de los vehículos es un problema que surge en muchos contextos. Cuestiones tales como programar el equipo de remoción de nieve, las entregas postales y bancarias, y los fletes a los clientes con requerimientos variables para diferentes lugares son algunos ejemplos. La programación de vehículos se estudia en la sección 6.6.
5. *Programación de vendedores*. Para compañías con sistemas de producción justo a tiempo, programar las entregas de los vendedores es un aspecto logístico importante. Las compras deben coordinarse con el sistema completo de entrega de productos para asegurar que los sistemas de producción JIT funcionen eficientemente. Vollman y otros (1992, p. 191) estudian la aplicación de la programación de los vendedores al sistema JIT en Steelcase. (JIT se estudia en los capítulos 1 y 7 de este libro.)

6. *Programación del proyecto*. Un proyecto puede dividirse en un conjunto de tareas interrelacionadas. Aunque algunas pueden hacerse en forma concurrente, muchas no pueden iniciarse sino hasta que se terminen otras. Los proyectos complejos pueden incluir miles de tareas individuales que deben coordinarse para que el proyecto se termine a tiempo y dentro del presupuesto. La programación del proyecto es un componente importante de la función de planeación, que estudiamos con detalle en el capítulo 9.
7. *Programación dinámica contra estática*. La mayor parte de la teoría de programación que revisamos en este capítulo considera el problema de programación como estático. Numerosos trabajos llegan simultáneamente para ser procesados en un conjunto de máquinas. En la práctica, muchos problemas de programación son dinámicos en el sentido de que los trabajos llegan continuamente fuera del tiempo estipulado. Un ejemplo es el problema que encara un controlador de tráfico aéreo, quien debe asignar las pistas de despegue y de aterrizaje para los aeroplanos que llegan. El problema es de tipo dinámico porque los aeroplanos llegan de manera aleatoria, y las pistas se liberan y se asignan también aleatoriamente. Los problemas de programación dinámica, estudiados en la sección 9 de este capítulo, se analizan usando las herramientas de la teoría de colas (estudiada con detalle en el suplemento 2, que sigue a este capítulo).

La programación es una función operativa compleja pero muy importante. El propósito de este capítulo es brindar al lector una idea sobre los tipos de resultados que pueden obtenerse al usar modelos analíticos, y mostrar cómo estos modelos pueden usarse para resolver ciertas clases de problemas de programación. Nuestro enfoque se concentra en la programación del taller de trabajos, aunque también consideramos algunos otros problemas de programación.

8.1 PROGRAMACIÓN DE LA PRODUCCIÓN Y JERARQUÍA DE LAS DECISIONES DE PRODUCCIÓN

Programar detalladamente los diferentes aspectos de la función de producción es crucial para controlar las operaciones de producción. Podemos considerar a la función de producción de una compañía como un proceso jerárquico. Primero, la compañía debe pronosticar la demanda de las ventas agregadas para algún horizonte de planeación previamente determinado. Estos pronósticos suministran la entrada para determinar la producción agregada y los niveles de la fuerza de trabajo para el horizonte de planeación. Las técnicas de planeación agregada se estudian en el capítulo 3. Entonces, el plan de producción agregada se traslada al programa maestro de producción (MPS). El MPS conduce a metas específicas de producción por producto e intervalo de tiempo.

La planeación de requerimientos de materiales (MRP), estudiada con detalle en el capítulo 7, es un método generado por el MPS para satisfacer metas específicas de producción del inventario de bienes terminados. El sistema MRP "explosiona" los niveles de producción resultantes tras el análisis de MPS en forma retrospectiva, con miras a obtener objetivos de producción para cada nivel de ensamblado por intervalo de tiempo. El resultado del análisis de MRP son las órdenes específicas de liberación planeadas específicas para los productos finales, los subensamblados y los componentes.

Finalmente, las órdenes de liberación planeadas deben trasladarse a un conjunto de tareas, así como también los datos asociados con éstas. Este nivel de planeación detallada conduce al programa de piso. Ya que el MRP u otro sistema de programación de lotes generalmente recomiendan las revisiones de las órdenes de liberación planeadas, los programas de piso cambian frecuentemente. En la figura 8-1 se muestra esquemáticamente la jerarquía de las decisiones de planeación de la producción.

El control de piso conlleva programar personal y equipo en un taller para satisfacer las fechas de vencimiento de un grupo de trabajos. Con frecuencia los trabajos deben procesarse en las máquinas del taller, según cierto orden o una secuencia únicos. La figura 8-2 muestra la disposición de un taller típico.

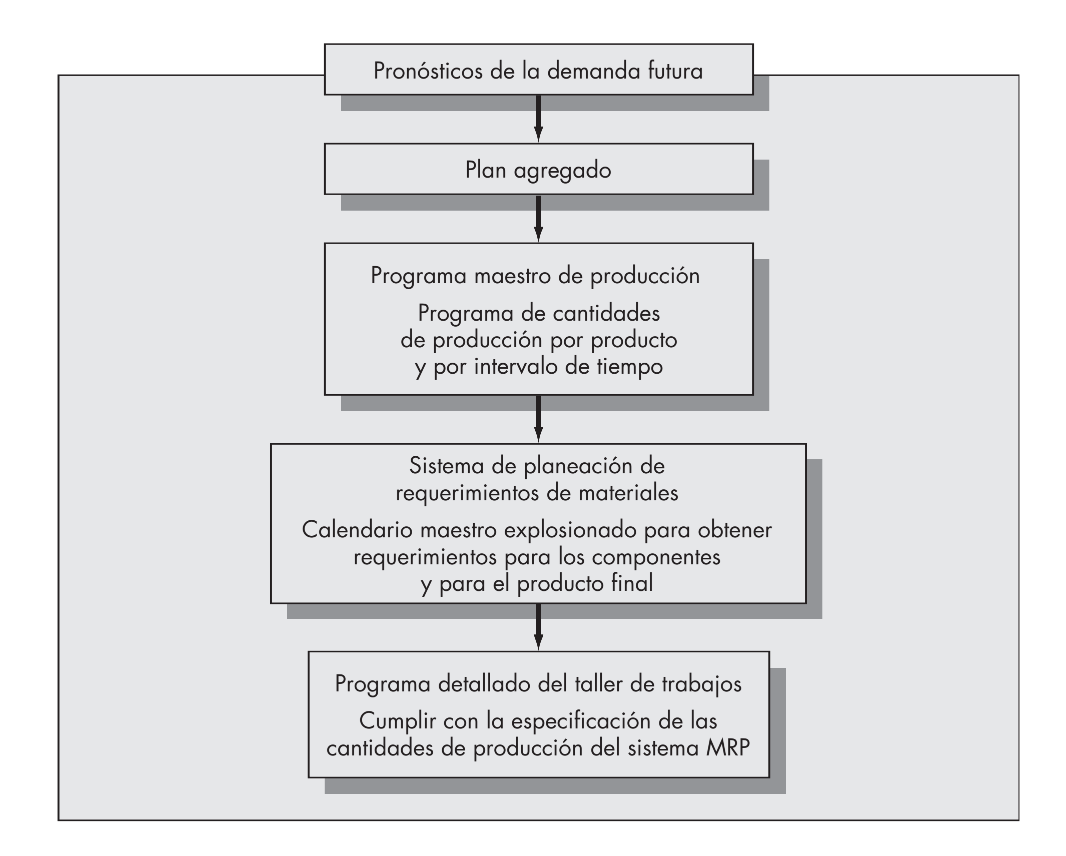


FIGURA 8-1
Jerarquía de las decisiones de producción.

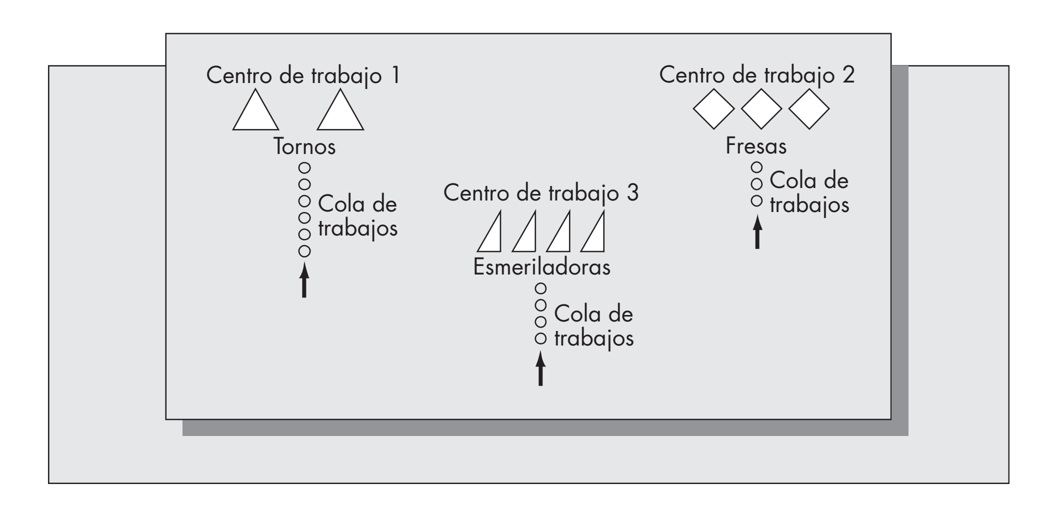


FIGURA 8-2
Disposición típica de un taller.

Tanto los trabajos como las máquinas se tratan de manera indivisible. Cuando las máquinas están ocupadas, los trabajos deben esperar, o hacer cola, para su procesamiento. A esto se le llama procesamiento discreto. La programación de la producción en industrias de proceso continuo, como la azucarera o las refinerías petroleras, tiene características muy diferentes.

Aunque hay muchos problemas asociados con la programación de operaciones, nuestra preocupación en este capítulo será la secuencia de los trabajos. Tenemos un conjunto de trabajos en espera de ser procesados en un grupo de máquinas, y el problema reside en cómo ordenarlos para optimizar algún criterio específico. La selección apropiada de la regla de secuenciación puede mejorar dramáticamente la tasa total del taller.

8.2 CARACTERÍSTICAS IMPORTANTES DE LOS PROBLEMAS DE PROGRAMACIÓN EN TALLERES

Los siguientes son aspectos importantes al determinar reglas de programación aproximadamente óptimas:

1. *El patrón de llegada de los trabajos*. Frecuentemente consideramos al problema del taller como estático: tomamos una "fotografía" del sistema en un punto del tiempo y procedemos a resolver el problema basándonos en el valor del estado presente. Sin embargo, el número de trabajos en espera de ser procesados cambia constantemente. Entonces, aunque muchos de los algoritmos de solución que consideramos juzgan al problema como estático, la mayoría de los problemas prácticos de programación de trabajos son dinámicos por naturaleza.
2. *Número y variedad de máquinas en el taller*. Un taller específico puede tener características únicas que podrían dificultar la implementación de una solución obtenida a partir de un algoritmo de programación. Por ejemplo, generalmente se supone que todas las máquinas de un tipo dado son idénticas. Sin embargo, éste no siempre es el caso. La tasa de producción de una máquina específica podría depender de varios factores, tales como las condiciones de la máquina o la pericia del operador. Según la disposición del taller y la naturaleza de los trabajos, podrían existir restricciones que harían irrealizables las soluciones obtenidas a partir de un procedimiento para "todo propósito".
3. *Número de trabajadores en el taller*. Tanto el número de trabajadores como la cantidad y variedad de máquinas en el taller determinan su capacidad. La planeación de ésta es un aspecto importante de la planeación de la producción. Muchos sistemas de control, tales como el MRP tradicional estudiado en el capítulo 7, no incorporan explícitamente las consideraciones de capacidad. Más aún, la capacidad es dinámica. La descompostura de una sola máquina o la pérdida de un empleado importante podrían conducir a un cuello de botella y a una reducción de la capacidad en el taller.
4. *Patrones específicos de flujo*. Las soluciones obtenidas de los algoritmos de programación que van a presentarse en este capítulo requieren que los trabajos se terminen en un orden fijo. Sin embargo, cada secuencia de trabajos a través de las máquinas conducen a un patrón de flujo de materiales a través del sistema. Debido a que con frecuencia los aspectos del manejo de materiales se tratan por separado de los aspectos de programación, pueden surgir patrones de flujo imprácticos.
5. *Evaluación de las reglas alternas*. La selección de los objetivos va a determinar qué tan adecuada y efectiva es una regla de secuenciación. Es común que más de un objetivo sea importante, por lo que es imposible determinar una regla óptima única. Por ejemplo, podría desearse minimizar el tiempo requerido para terminar todos los trabajos, pero también podría desearse limitar el retraso máximo de cualquier trabajo individual.

Objetivos de la administración de trabajo en un taller

Una de las dificultades de la programación es que con frecuencia se presentan metas conflictivas. Los objetivos de las diferentes partes de la compañía no siempre son los mismos. Algunos de los propósitos más comunes son

1. Cumplir con las fechas de entrega.
2. Minimizar el inventario del trabajo en proceso (WIP).
3. Minimizar el tiempo promedio de flujo a través del sistema.
4. Suministrar un elevado tiempo de uso de máquina/trabajador. (Minimizar el tiempo muerto de máquina/trabajador.)
5. Suministrar información exacta del estado de los trabajos.
6. Reducir los tiempos de preparación.
7. Minimizar los costos de producción y de los trabajadores.

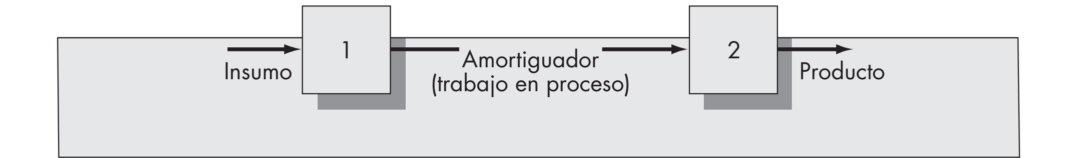


FIGURA 8-3
Un proceso compuesto de dos operaciones en serie.

Obviamente es imposible optimizar simultáneamente los siete objetivos. En especial (1) y (3), que principalmente se abocan a ofrecer un alto nivel de servicio al consumidor, y (2), (4), (6) y (7), dirigidos a suministrar un alto nivel de eficiencia en planta. Determinar el intercambio entre costo y calidad es uno de los aspectos estratégicos más importantes que actualmente enfrenta una compañía.

Algunos de estos objetivos entran en conflicto. Si la finalidad principal consiste en reducir el inventario del trabajo en proceso (como por ejemplo, con los sistemas de control de inventario justo a tiempo, estudiados en el capítulo 7), es probable que aumente el tiempo muerto del trabajador. A medida que el sistema se estrecha por la reducción del inventario dentro y entre las operaciones de manufactura, las diferencias en la tasa de producción de una parte del sistema a otra pueden forzar la espera de operaciones más rápidas. Aun cuando los partidarios de la filosofía JIT no lo recomiendan, los inventarios de amortiguamiento entre las operaciones pueden reducir significativamente el tiempo muerto.

Como ejemplo, considere el sistema simple compuesto de dos operaciones en serie, ilustrado en la figura 8-3. Si el inventario del trabajo en proceso es cero, entonces la producción del sistema para cualquier instante de tiempo está gobernada por la producción más pequeña de las dos operaciones. Si la operación 1 se detiene temporalmente por una descompostura de la máquina, entonces la operación 2 también debe quedar parada. Sin embargo, si hay un inventario de amortiguamiento colocado entre las operaciones, entonces 2 puede continuar operando mientras 1 se somete a una reparación o una recalibración.

Encontrar la combinación apropiada entre el inventario WIP y el tiempo muerto del trabajador equivale a escoger un punto en la curva de intercambios de estos objetivos conflictivos. (Las curvas de intercambio se estudiaron en el capítulo 5 en el contexto del control de inventario de artículos múltiples.) En la figura 8-4*a* se ilustra una curva de este tipo. Un movimiento de un punto a otro a lo largo de una curva de intercambio no necesariamente implica que el sistema ha mejorado, sino más bien que se están aplicando pesos diferentes a los dos objetivos. Una verdadera mejora en el sistema general implicaría que la curva de intercambio completa experimentara un desplazamiento hacia abajo, tal como el que se ilustra en la figura 8-4*b*.

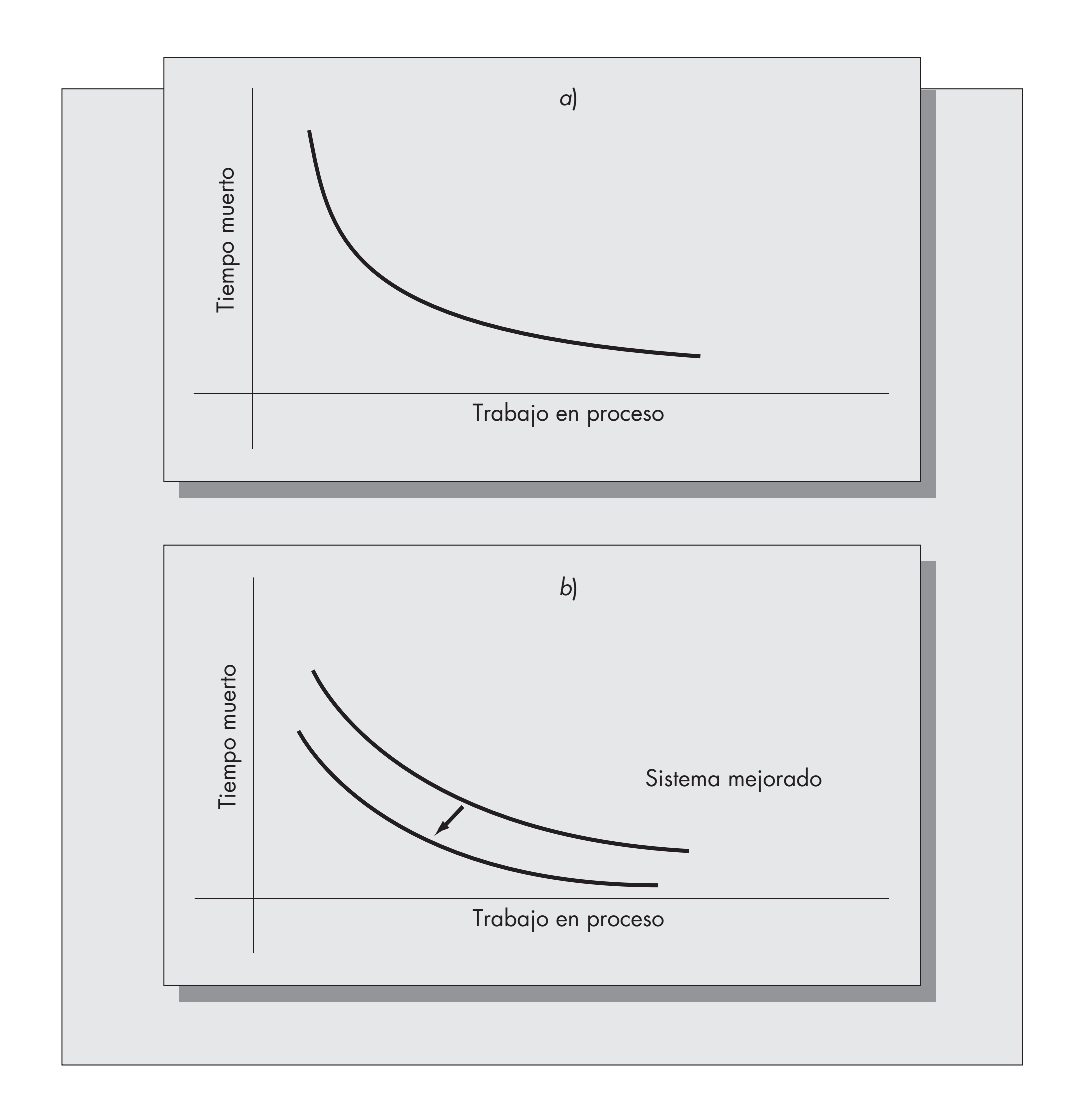
8.3 TERMINOLOGÍA DE LA PROGRAMACIÓN DE TRABAJOS POR TALLER

En general, un problema que se presenta en la programación de trabajos en un taller es que los n trabajos deben procesarse a través de m máquinas. La complejidad del problema depende de varios factores, tales como las secuencias de trabajo permisibles y qué criterios de optimización se escogen. En esta sección definimos algunos de los términos que se usarán a lo largo del capítulo.

1. *Taller de flujo*. En un taller de flujo cada uno de los n trabajos debe procesarse a través de las m máquinas en el mismo orden, y cada trabajo se procesa exactamente una vez en cada máquina. Esto es lo que comúnmente entendemos por una línea de ensamblado.

2. *Taller*. Un taller generalmente difiere de un taller de flujo, en que no se considera que todos los trabajos necesiten exactamente m operaciones, y algunos trabajos pueden requerir operaciones múltiples en una sola máquina. Aún más, en un taller cada trabajo puede tener una secuencia de operaciones diferente. Los problemas del taller en general son muy complejos, y no se dispone de algoritmos de solución de uso múltiple.

FIGURA 8-4
Objetivos conflictivos en la administración de un taller.



3. *Procesamiento paralelo contra procesamiento secuencial*. La mayoría de los problemas que vamos a considerar incluyen el procesamiento secuencial. Esto significa que las m máquinas son diferentes, y que éstas realizan varias operaciones distintas.. En el procesamiento paralelo suponemos que las máquinas son idénticas, y que cualquier trabajo puede ser procesado en cualquier máquina. Un ejemplo de procesamiento paralelo es una central telefónica, en la que las llamadas se procesan a través del siguiente servidor disponible. El procesamiento paralelo se estudia en el contexto de la programación estocástica en la sección 8.8.

4. *Tiempo de flujo*. El tiempo de flujo del trabajo i es el tiempo que transcurre desde el inicio del primer trabajo en la primera máquina hasta la terminación del trabajo i. En forma equivalente, es el lapso que el trabajo i reside en el sistema. El *tiempo de flujo medio*, que es una medida común del rendimiento de un sistema, es el promedio aritmético de los tiempos de flujo para los n trabajos.

5. *Terminación*. Es el tiempo de flujo del último trabajo terminado. Es también el tiempo requerido para terminar los n trabajos.

6. *Retardos y retrasos*. El retardo es la diferencia positiva entre el tiempo de terminación (tiempo de flujo) y la fecha de vencimiento de un trabajo. Un trabajo retardado es aquel que se termina después de su fecha de vencimiento. El retraso se refiere a la diferencia entre el tiempo

de terminación del trabajo y su fecha de vencimiento, y difiere del retardo en que el retraso puede ser positivo o negativo. La minimización del retardo promedio y del retardo máximo es también un objetivo común de la programación.

8.4 UNA COMPARACIÓN DE REGLAS DE SECUENCIA ESPECÍFICA

En la comparación y la evaluación de las reglas de secuenciación, consideramos al taller en un instante fijo en el tiempo. Esta sección va a centrarse solamente en una máquina. Suponga que existe un grupo de trabajos que deben procesarse en ésta y que cada trabajo tiene asociado un tiempo de procesamiento y una fecha de vencimiento. Comparamos el desempeño de cuatro reglas de secuenciación que comúnmente se usan en la práctica. El objetivo de esta sección es ilustrar cómo afectan estas reglas de secuenciación a las diferentes medidas del desempeño del sistema.

Comparamos las siguientes cuatro reglas de secuenciación:

1. *Primeras llegadas, primeras salidas* (FCFS). Los trabajos se procesan en el orden en que entraron al taller.
2. *Tiempo de procesamiento más corto primero* (SPT). Los trabajos se ordenan de manera ascendente correspondiendo con los tiempos de procesamiento. El trabajo con el tiempo de procesamiento más corto es el primero, el trabajo con el siguiente tiempo de procesamiento más corto es el segundo, etcétera.
3. *Primera fecha de entrega* (EDD). Los trabajos se ordenan de manera ascendente de acuerdo con sus fechas de vencimiento. El trabajo con la fecha de entrega más próxima es el primero, el trabajo con la siguiente fecha de entrega más próxima es el segundo, etcétera.
4. *Programación basada en la razón crítica* (CR). En este caso se requiere calcular la relación del tiempo de procesamiento del trabajo, dividido entre el tiempo restante hasta la fecha de vencimiento, y programar el trabajo con la siguiente relación más grande.

Comparamos el rendimiento de estas cuatro reglas para un caso específico basado en el tiempo de flujo medio, el retardo promedio y el número de trabajos retardados. El objetivo del siguiente ejemplo es ayudar al lector a desarrollar una intuición de la mecánica de programación antes de presentar resultados formales.

Ejemplo *8.1* Un centro de maquinado en un taller donde se realizan trabajos para una compañía local de fabricación tiene cinco trabajos no procesados que se quedan pendientes para un instante específico en el tiempo. Los trabajos están rotulados como 1, 2, 3, 4 y 5 en el orden de ingreso al taller. Los tiempos de procesamiento y las fechas de entrega respectivos están dados en la siguiente tabla.

Número de trabajo	Tiempo de procesamiento	Fecha de entrega
1	11	61
2	29	45
3	31	31
4	1	33
5	2	32

Primeras llegadas, primeras salidas

Como los trabajos ingresaron al taller en la secuencia en que fueron numerados, la programación FCFS implica que los trabajos están ordenados de la siguiente manera: 1, 2, 3, 4, 5.

Esto conduce a

Secuencia	Tiempo de terminación	Fecha de entrega	Retardo
1	11	61	0
2	40	45	0
3	71	31	40
4	72	33	39
5	74	32	42
Totales	268		121

$$\text{Tiempo de flujo medio} = 268/5 = 53.6.$$
$$\text{Retardo promedio} = 121/5 = 24.2.$$
$$\text{Número de trabajos retardados} = 3.$$

El retardo de un trabajo es igual a cero si el trabajo se termina antes de su fecha de entrega y es igual al número de días de retraso si el trabajo se termina después de su fecha de entrega.

Tiempo de procesamiento más corto

Aquí los trabajos son ordenados de acuerdo con el tiempo de procesamiento creciente.

Trabajo	Tiempo de procesamiento	Tiempo de terminación	Fecha de entrega	Retardo
4	1	1	33	0
5	2	3	32	0
1	11	14	61	0
2	29	43	45	0
3	31	74	31	43
Totales		135		43

$$\text{Tiempo de flujo medio} = 135/5 = 27.0.$$
$$\text{Retardo promedio} = 43/5 = 8.6.$$
$$\text{Número de trabajos retardados} = 1.$$

Primera fecha de entrega

Aquí los trabajos se terminan en el orden de sus fechas de entrega.

Trabajo	Tiempo de procesamiento	Tiempo de terminación	Fecha de entrega	Retardo
3	31	31	31	0
5	2	33	32	1
4	1	34	33	1
2	29	63	45	18
1	11	74	61	13
Totales		235		33

$$\text{Tiempo de flujo medio} = 235/5 = 47.0.$$
$$\text{Retardo promedio} = 33/5 = 6.6.$$
$$\text{Número de trabajos retardados} = 4.$$

Programación basada en la razón crítica

Después del procesamiento de cada trabajo, calculamos

$$\frac{\text{Fecha de entrega} - \text{Tiempo presente}}{\text{Tiempo de procesamiento}},$$

que se conoce como la razón crítica, y programamos el siguiente trabajo con objeto de minimizar el valor de la razón crítica. La idea que da sustento a la programación por la razón crítica es suministrar un balance entre SPT, que solamente considera el tiempo de procesamiento, y EDD, que solamente considera las fechas de entrega. La razón se hará más pequeña a medida que el tiempo presente se aproxima a la fecha de entrega, y se dará más prioridad a aquellos trabajos con tiempos de procesamiento mayores. Una desventaja del método es que las razones críticas necesitan recalcularse cada vez que se programe un trabajo.

Es posible que el numerador sea negativo para alguno o para todos los trabajos restantes. Si esto ocurre significa que el trabajo está retrasado, y vamos a suponer que los trabajos retrasados se programen automáticamente como los siguientes. Si hay más de un trabajo retrasado, entonces los trabajos retrasados se programen de acuerdo a la secuencia SPT.

Primero calculamos las razones críticas iniciando en el instante $t = 0$.

Tiempo presente: $t = 0$

Trabajo	Tiempo de procesamiento	Fecha de entrega	Razón crítica
1	11	61	61/11 (5.545)
2	29	45	45/29 (1.552)
3	31	31	31/31 (1.000)
4	1	33	33/1 (33.00)
5	2	32	32/2 (16.00)

El valor mínimo corresponde al trabajo 3, de modo que el trabajo 3 se ejecuta primero. Como el trabajo 3 requiere 31 unidades de tiempo para procesarse, debemos actualizar todas las razones críticas con objeto de determinar el siguiente trabajo que va a procesarse. Ponemos el reloj en el instante $t = 31$ y recalculamos las razones críticas.

Tiempo presente: $t = 31$

Trabajo	Tiempo de procesamiento	Fecha de entrega − tiempo presente	Razón crítica
1	11	30	30/11 (2.727)
2	29	14	14/29 (0.483)
4	1	2	2/1 (2.000)
5	2	1	1/2 (0.500)

El mínimo es 0.483, que corresponde al trabajo 2. Así, el trabajo 2 se programa enseguida. Como el trabajo 2 tiene un tiempo de procesamiento de 29, actualizamos el reloj para el instante $t = 31 + 29 = 60$.

Tiempo presente: $t = 60$

Trabajo	Tiempo de procesamiento	Fecha de entrega − tiempo presente	Razón crítica
1	11	1	1/11 (.0909)
4	1	−27	$-27/1 < 0$
5	2	−28	$-28/2 < 0$

Ahora los trabajos 4 y 5 están retardados, de modo que se les da prioridad y se programan enseguida. Como se programan en el orden SPT, se ejecutan en la secuencia trabajo 4, luego trabajo 5. Finalmente, el trabajo 1 se programa al último.

Resumen de los resultados de la programación de la razón crítica

Trabajo	Tiempo de procesamiento	Tiempo de terminación	Retardo
3	31	31	0
2	29	60	15
4	1	61	28
5	2	63	31
1	11	74	13
Totales		289	87

$$\text{Tiempo de flujo medio} = 289/5 = 57.8.$$

$$\text{Retardo promedio} = 87/5 = 17.4.$$

$$\text{Número de trabajos retardados} = 4.$$

Resumimos los resultados de esta sección para las cuatro reglas de programación:

Resumen de los resultados de cuatro reglas de programación

Regla	Tiempo de flujo medio	Retardo promedio	Número de trabajos retardados
FCFS	53.6	24.2	3
SPT	27.0	8.6	1
EDD	47.0	6.6	4
CR	57.8	17.4	4

8.5 OBJETIVOS DE LA ADMINISTRACIÓN DE TRABAJO EN TALLER: UN EJEMPLO

Ejemplo *8.2*

Un controlador de tráfico aéreo se enfrenta al problema de programar el aterrizaje de cinco aeronaves. Basándose en la posición y en los requerimientos de pista de cada aeroplano, estima los siguientes tiempos de aterrizaje:

Aeroplano:	1	2	3	4	5
Tiempo (en minutos):	26	11	19	16	23

Solamente un avión puede aterrizar a la vez. El problema es esencialmente el mismo que el de programar cinco trabajos para una sola máquina. Los aeroplanos corresponden a los trabajos, los tiempos de aterrizaje a los tiempos de procesamiento, y la pista a la máquina.

1. Con la información dada, dos objetivos razonables serían minimizar el tiempo total requerido (es decir, la terminación) o el tiempo promedio requerido (el tiempo de flujo medio) para que todos los aeroplanos aterricen. Es claro que la terminación de cualquier secuencia es de 95 minutos, la suma de los tiempos de aterrizaje. Sin embargo, como vimos en el ejemplo 8.1, el tiempo de flujo medio no es independiente de la secuencia y la regla del tiempo de procesamiento más corto minimiza al tiempo de flujo medio. En la sección 8.6 mostraremos que SPT es la regla óptima de secuenciación para minimizar el tiempo de flujo medio para una sola máquina en general.

2. Un objetivo alternativo puede ser aterrizar la mayor cantidad de gente tan rápido como sea posible. En este caso también necesitaríamos conocer el número de pasajeros que viaja en cada aeroplano.

Suponga que estos números son los siguientes:

Aeroplano	1	2	3	4	5
Tiempo de aterrizaje	26	11	19	16	23
Número de pasajeros	180	12	45	75	252

El objetivo apropiado en este caso podría ser minimizar la terminación ponderada o la suma ponderada de los tiempos de terminación, donde los pesos corresponderían al número de pasajeros en cada aeroplano. Observe ahora que la función objetivo tendría unidades de pasajero-minutos.

3. Un aspecto que todavía no tratamos es el tiempo de programación de la llegada de cada aeroplano. Suponga los siguientes datos:

Aeroplano	1	2	3	4	5
Tiempo de aterrizaje	26	11	19	16	23
Hora programada de llegada	5:30	5:45	5:15	6:00	5:40

Las reglas de secuenciación que ignoran las fechas de entrega podrían dar resultados muy pobres en cuanto a satisfacer las horas de llegada. Algunos objetivos posibles relacionados con las fechas de entrega incluyen la minimización del retardo promedio y la minimización del retardo máximo.

4. Hasta ahora hemos ignorado condiciones especiales que favorecen a algunos aeroplanos sobre otros. Suponga que el aeroplano número 4 tiene un nivel de combustible críticamente bajo. Esto probablemente conduciría a que éste tomara precedencia. También podrían surgir restricciones de prioridad de otras maneras: es posible dar prioridad a los aeroplanos que están programados para continuar los vuelos o a los aeroplanos que transportan carga valiosa o perecedera.

El propósito de esta sección fue demostrar las dificultades para seleccionar una función objetivo para los problemas de secuenciación de trabajos. La secuencia óptima es altamente sensible a la selección del objetivo, y el objetivo apropiado no siempre es obvio.

Problemas para las secciones 8.1 a 8.5

1. Discuta los siguientes objetivos listados y la relación que tiene cada uno con el desempeño de un taller.

   a) Reducir el inventario WIP.

   b) Suministrar un alto nivel de servicio al cliente.

   c) Reducir el tiempo muerto del trabajador.

   d) Mejorar el rendimiento de la fábrica.

2. En el problema 1, ¿por qué *a* y *c* son objetivos conflictivos, y por qué *b* y *d* también lo son?
3. Defina los siguientes términos:

   a) Taller de flujo.

   b) Taller.

   c) Procesamiento secuencial contra procesamiento paralelo.

   d) Terminación.

   e) Retardo.

4. Cuatro camiones, 1, 2, 3, y 4, están esperando en una rampa de carga en la compañía XYZ que tiene solamente una crujía de servicio. Los camiones están rotulados en el orden de llegada a la rampa. Suponga que la hora presente es 1:00 p. m. En la siguiente tabla se dan

los tiempos de descarga requeridos por cada camión y los tiempos de entrega del material en la planta.

Camión	Tiempo de descarga (minutos)	Hora de entrega del material
1	20	1:25 p.m.
2	14	1:45 p.m.
3	35	1:50 p.m.
4	10	1:30 p.m.

Determine el programa que resulta para cada una de las reglas FCFS, SPT, EDD, y CR. En cada caso calcule el tiempo de flujo medio, el retardo promedio y el número de trabajos retardados.

5. Cinco trabajos deben programarse para el procesamiento por lotes en un sistema de computadora *mainframe*. Los tiempos de procesamiento y la hora prometida para cada uno de los trabajos se listan aquí.

Trabajo	1	2	3	4	5
Tiempo de procesamiento	40 min	2.5 h	20 min	4 h	1.5 h
Hora prometida	11:00 a.m.	2:00 p.m.	2:00 p.m.	1:00 p.m.	4:00 p.m.

Suponga que la hora presente es 10:00 a. m.

a) Si los trabajos están programados de acuerdo con SPT, encuentre el retardo de cada trabajo y el retardo promedio de todos los trabajos.

b) Repita el cálculo de la parte *a* para la programación EDD.

8.6 UNA INTRODUCCIÓN A LA TEORÍA DE SECUENCIACIÓN PARA UNA SOLA MÁQUINA

Suponga que n trabajos van a procesarse a través de una máquina. Para cada trabajo i, defina las siguientes cantidades:

t_i = Tiempo de procesamiento para el trabajo i,

d_i = Fecha de entrega del trabajo i,

W_i = Tiempo de espera para el trabajo i,

F_i = Tiempo de flujo para el trabajo i,

L_i = Retraso del trabajo i,

T_i = Retardo del trabajo i,

E_i = Anticipación del trabajo i.

El tiempo de procesamiento y la fecha de entrega son constantes que están incorporadas a la descripción de cada trabajo. El tiempo de espera de un trabajo es el que debe esperar antes de que pueda comenzar su procesamiento. Para los casos que consideramos, es también la suma de los tiempos de procesamiento de todos los trabajos anteriores. El tiempo de flujo es simplemente el tiempo de espera más el de procesamiento del trabajo ($F_i = W_i + t_i$). El tiempo de flujo del trabajo i y el de terminación del trabajo i son los mismos. Definiremos el retraso del trabajo i como $L_i = F_i - d_i$, y supondremos que el retraso puede ser una cantidad positiva o bien, una negativa. El retardo es la parte positiva del retraso ($T_i = \text{máx}[L_i, 0]$), y la anticipación es la parte negativa del retraso ($E_i = \text{máx}[-L_i, 0]$).

Otras cantidades relacionadas son el retardo máximo $T_{\text{máx}}$, dado por la fórmula

$$T_{\text{máx}} = \text{máx}\{T_1, T_2, \ldots, T_n\},$$

y el tiempo de flujo medio F', dado por la fórmula

$$F' = \frac{1}{n}\sum_{i=1}^{n} F_i.$$

Como solamente estamos considerando una sola máquina, cada calendario puede representarse por una permutación (es decir, ordenamiento) de los enteros 1, 2, . . . , n. Hay exactamente $n!$ calendarios de permutación diferentes [$n! = n(n-1)\cdots(2)(1)$].

Programación del tiempo de procesamiento más corto

Tenemos el siguiente resultado:

Teorema 8.1 **La regla de programación que minimiza el tiempo de flujo medio F' es SPT.**

El teorema 8.1 es fácil de probar. Sea [1], [2], . . . , [n] cualquier permutación de los enteros 1, 2, 3, . . . , n. El tiempo de flujo del trabajo que se programa en la posición k está dado por

$$F_{[k]} = \sum_{i=1}^{k} t_{[i]}.$$

Se concluye que el tiempo de flujo medio esta dado por

$$F' = \frac{1}{n}\sum_{k=1}^{n} F_{[k]} = \frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{k} t_{[i]}.$$

El término de doble sumatoria puede escribirse de forma diferente. Al expandir la doble sumatoria, obtenemos

$$\begin{aligned} k &= 1\text{: } t_{[1]} \\ k &= 2\text{: } t_{[1]} + t_{[2]} \\ &\vdots \\ k &= n\text{: } t_{[1]} + t_{[2]} + \cdots + t_{[n]}. \end{aligned}$$

Al sumar por columnas en vez de sumar por renglones, podemos reescribir F' en la forma

$$nt_{[1]} + (n-1)t_{[2]} + \cdots + t_{[n]},$$

que claramente se minimiza al establecer

$$t_{[1]} \le t_{[2]} \le \cdots \le t_{[n]},$$

que es exactamente la regla de secuenciación SPT.

Tenemos el siguiente corolario del teorema 8.1.

Corolario 8.1 **Las siguientes medidas son equivalentes:**

1. **Tiempo de flujo medio**
2. **Tiempo de espera medio**
3. **Retraso medio**

En conjunto, el corolario 8.1 y el teorema 8.1 establecen que SPT minimiza el tiempo de flujo medio, el tiempo de espera medio y el retraso medio para la secuenciación de una sola máquina.

Programación de la primera fecha de entrega

Si el objetivo es minimizar el retraso máximo, entonces los trabajos deben ordenarse de acuerdo con sus fechas de entrega. Es decir, $d_{[1]} \leq d_{[2]} \leq \cdots \leq d_{[n]}$. No presentaremos una prueba de este resultado. La idea que fundamenta la prueba es seleccionar algún programa que no ordene los trabajos en relación con sus fechas de entrega; eso implica que existe un valor k tal que $d_{[k]} > d_{[k+1]}$. Se demuestra que al intercambiar las posiciones de los trabajos k y $k + 1$, se reduce el retraso máximo.

Minimización del número de trabajos retrasados

Hay muchos ejemplos en los cuales la penalización por un trabajo atrasado (retardado) permanece igual sin importar qué tan grande sea el retraso. Por ejemplo, cualquier atraso en la terminación de todas las tareas que se requieren para la preparación de un lanzamiento al espacio podría causar el aborto del lanzamiento, independientemente de la magnitud del atraso.

Vamos a describir el algoritmo de Moore (1968), que minimiza el número de trabajos retardados para el problema de una sola máquina.

Paso 1. Ordene los trabajos de acuerdo con la fecha de la primera entrega para obtener la solución inicial. Es decir $d_{[1]} \leq d_{[2]} \leq \cdots \leq d_{[n]}$.

Paso 2. Encuentre el primer trabajo retardado en la secuencia presente, digamos el trabajo $[i]$. Si no existe ninguno, vaya al paso 4.

Paso 3. Considere los trabajos [1], [2], . . . , $[i]$. Rechace el trabajo con el mayor tiempo de procesamiento. Regrese al paso 2.

Paso 4. Forme una secuencia óptima tomando la secuencia presente y adjuntándole los trabajos rechazados. Los trabajos adjuntos a la secuencia presente pueden programarse en cualquier orden porque constituyen los trabajos retardados.

Ejemplo *8.3*

Un taller de máquinas procesa las órdenes a la medida de varios clientes. Una de las máquinas, una esmeriladora, tiene seis trabajos en espera de ser ejecutados. Aquí se dan los tiempos de procesamiento y las fechas prometidas de entrega (ambos en horas) para los seis trabajos.

Trabajo	1	2	3	4	5	6
Fecha de entrega	15	6	9	23	20	30
Tiempo de procesamiento	10	3	4	8	10	6

El primer paso es ordenar los trabajos de acuerdo con la regla EDD.

Trabajo	2	3	1	5	4	6
Fecha de entrega	6	9	15	20	23	30
Tiempo de procesamiento	3	4	10	10	8	6
Tiempo de terminación	3	7	17	27	35	41

Vemos que el primer trabajo retardado es el trabajo 1, y que hay un total de cuatro trabajos retardados. Ahora consideramos los trabajos 2, 3, y 1, y rechazamos el trabajo con el mayor tiempo de procesamiento. Se trata, evidentemente, del trabajo 1. En este momento, la nueva secuencia presente es

Trabajo	2	3	5	4	6
Fecha de entrega	6	9	20	23	30
Tiempo de procesamiento	3	4	10	8	6
Tiempo de terminación	3	7	17	25	31

El primer trabajo retardado en la secuencia presente es ahora el trabajo 4. Consideramos la secuencia 2, 3, 5, 4, y rechazamos el trabajo con el mayor tiempo de procesamiento, que es el trabajo 5. La secuencia presente es ahora

Trabajo	2	3	4	6
Fecha de entrega	6	9	23	30
Tiempo de procesamiento	3	4	8	6
Tiempo de terminación	3	7	15	21

Es evidente que en esta etapa no hay trabajos retardados. La secuencia óptima es 2, 3, 4, 6, 5, 1, o bien, 2, 3, 4, 6, 1, 5. En cualquiera de los dos casos el número de trabajos retardados es exactamente 2.

Restricciones de precedentes: Algoritmo de Lawler

El algoritmo de Lawler (Lawler, 1973) es una técnica poderosa para la solución de varios problemas restringidos de programación. Se supone que la función objetivo es de la forma

$$\text{mín} \max_{1 \leq i \leq n} g_i(F_i)$$

donde g_i es cualquier función no decreciente del tiempo de flujo F_i. Además, el algoritmo maneja *cualesquiera* restricciones de precedencia. Se presentan restricciones de precedencia cuando ciertos trabajos deben terminarse antes de que otros trabajos puedan comenzar; son muy comunes en los problemas de programación. Algunos ejemplos de funciones g_i que pueden considerarse son $g_i(F_i) = F_i - d_i = L_i$, que corresponde a minimizar el retraso máximo, o sea $g_i(F_i) = \text{máx}(F_i - d_i, 0)$, que equivale a la minimización del retraso máximo.

El Algoritmo

El algoritmo de Lawler programa primero el trabajo que debe terminarse en último lugar, luego el siguiente trabajo que debe terminarse después del último, etc. En cada etapa se determina el conjunto de trabajos que no se requiere que precedan a ningún otro. Denominemos a este conjunto V. Entre el conjunto V, seleccione el trabajo k que satisfaga

$$g_k(\tau) = \min_{i \in V} (g_i(\tau)),$$

donde $\tau = \sum_{i=1}^{n} t_i$ y corresponde al tiempo de procesamiento de la secuencia presente.

Ahora el trabajo k se programa como el último. Considere los trabajos restantes y determine nuevamente el conjunto de trabajos que no se requiere que precedan a ningún otro trabajo restante. Tras programar el trabajo k, este conjunto pudo haber cambiado. El valor de τ se reduce por t_k y ahora se determina el trabajo programado enseguida del último. El proceso se continúa hasta que se programan todos los trabajos. Observe que a medida que se programan los trabajos, pueden relajarse algunas de las restricciones de precedencia, de modo que es probable que el conjunto V cambie para cada iteración.

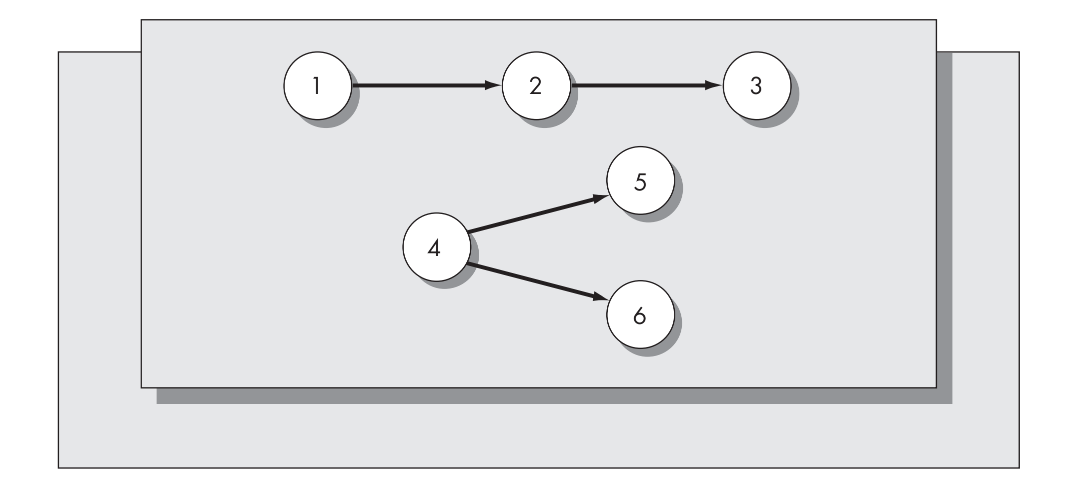
Ejemplo *8.4*

Tony D'Amato dirige un taller local de hojalatería que hace trabajos de hojalatería y pintura. Una cierta mañana de lunes tiene seis automóviles en espera de ser reparados. Tres (1, 2, y 3) son de una compañía de renta de automóviles y él está de acuerdo en terminar estos automóviles en el orden de las fechas en que se prometieron. Los automóviles 4, 5, y 6 son de un distribuidor minorista que ha solicitado que el automóvil 4 se termine primero porque lo está esperando un cliente. Las restricciones de precedencia resultantes pueden representarse como dos redes inconexas, según se ilustra en la figura 8-5.

Los tiempos requeridos para reparar cada uno de los automóviles (en días) y las fechas de terminación prometidas asociadas son

Trabajo	1	2	3	4	5	6
Tiempo de procesamiento	2	3	4	3	2	1
Fecha de entrega	3	6	9	7	11	7

FIGURA 8-5
Restricciones de precedencia para el ejemplo 8.4.



Determine como debe programarse la reparación de los automóviles en el taller con objeto de minimizar el retardo máximo.

Solución

1. Primero encontramos el último trabajo programado (el sexto). Entre los candidatos para la última posición están aquellos trabajos que no son predecesores de otros trabajos. Estos son 3, 5 y 6. El tiempo total de procesamiento de todos los trabajos es $2 + 3 + 4 + 3 + 2 + 1 = 15$. (Éste es el valor presente de τ.) Como el objetivo es minimizar el retardo máximo, comparamos los retardos de estos tres trabajos y escogemos el que tenga el menor valor. Obtenemos mín$\{15 - 9, 15 - 11, 15 - 7\}$ = mín$\{6, 4, 8\} = 4$, que corresponde al trabajo 5. Así que el trabajo 5 se programa como el último (posición 6).

2. Enseguida encontramos el trabajo programado en quinto lugar. Los candidatos son solamente los trabajos 3 y 6. En este momento el valor de τ es $15 - 2 = 13$. Así, encontramos mín$\{13 - 9, 13 - 7\}$ mín$\{4, 6\} = 4$, que corresponde al trabajo 3. Entonces, el trabajo 3 se programa en la quinta posición.

3. Encuentre el trabajo programado en cuarto lugar. Como el trabajo 3 ya no está en la lista, ahora el trabajo 2 se transforma en candidato. El valor presente de $\tau = 13 - 4 = 9$. Así, comparamos mín$\{9 - 6, 9 - 7\}$ = mín$\{3, 2\} = 2$, que corresponde al trabajo 6. Programe el trabajo 6 en la cuarta posición.

4. Encuentre el trabajo programado en tercer lugar. Se ha programado el trabajo 6, de modo que ahora el trabajo 4 se transforma en un candidato junto con el trabajo 2, y $\tau = 9 - 1 = 8$. Entonces, buscamos mín$\{8 - 6, 8 - 7\}$ = mín$\{2, 1\} = 1$, que se presenta en el trabajo 4.

5. En este momento encontraríamos al trabajo programado en segundo lugar. Sin embargo, nos quedan solamente los trabajos 1 y 2, los cuales, debido a las restricciones de precedencia, deben programarse en el orden 1-2.

Resumiendo los resultados, la secuencia óptima para reparar los automóviles es 1-2-4-6-3-5.

Con miras a determinar el valor de la función objetivo, el retardo máximo, calculamos el tiempo de flujo para cada trabajo y lo comparamos con la fecha de vencimiento. Tenemos

Trabajo	Tiempo de procesamiento	Tiempo de flujo	Fecha de entrega	Retardo
1	2	2	3	0
2	3	5	6	0
4	3	8	7	1
6	1	9	7	2
3	4	13	9	4
5	2	15	11	4

Así, el retardo máximo es de cuatro días. El lector debe convencerse de que cualquier otra secuencia conduce a un retardo máximo de cuando menos cuatro días.

Aplicación instantánea

COMPAÑIAS OPERADORAS DE AVIONES DE TIEMPO COMPARTIDO AHORRAN MILLONES CON EL SISTEMA DE PROGRAMACIÓN

Celebridades, ejecutivos, y deportistas profesionales conforman una buena parte del grupo que usa aviones privados. Para muchas de estas personas, comprar un aeroplano resultaría absurdo. Una opción atractiva es la propiedad fraccionaria, especialmente para aquellos que vuelan sólo ocasionalmente. La propiedad fraccionaria de los aeroplanos privados ofrece la flexibilidad de viajar a más de 5,000 destinos (en comparación con aproximadamente 500 de las líneas aéreas comerciales). Otras ventajas son la privacidad, el servicio personalizado, un número menor de demoras, además de la posibilidad de manejar sus negocios durante el viaje.

El concepto del programa fraccionario de aeronaves es similar al de un condominio de tiempo compartido, excepto que a los propietarios de la aeronave se les garantiza acceso en cualquier momento con solamente cuatro horas de anticipación. Las tarifas se basan en el número de horas de vuelo que va a requerir el propietario: si comparte un octavo se les asignan 100 horas de tiempo anual de vuelo, un cuarto 200 horas, etc. El sistema completo está coordinado por la compañía de manejo de fracciones (FMC: *fraccional management company*). Es evidente que el problema de programar los aviones y las tripulaciones puede ser muy complejo.

Al programar los aviones y las tripulaciones, el FMC necesita determinar horarios que (1) satisfagan las peticiones de los clientes con respecto al tiempo, (2) satisfagan las restricciones de mantenimiento y de la tripulación, y (3) permitan la asignación específica de los viajes de las aeronaves y las solicitudes. La capacidad de lucro de FMC dependerá de la eficiencia con que desempeñe estas tareas. Un grupo de consultores abordó este problema y desarrolló un sistema de programación conocido como ScheduleMiser.[1] Los insumos de éste son las solicitudes de viaje, la disponibilidad de las aeronaves, y las restricciones de las aeronaves para un horizonte de planeación específico. Observe que aun cuando se les garantice el servicio a los propietarios con solamente cuatro horas de anticipación, la gran mayoría de los viajes se reservan por lo menos con tres días de anticipación o más. Esto da un perfil confiable a la FMC de la demanda para un horizonte de planeación de dos a tres días. Observe que los horarios de las aeronaves deben coordinarse con los horarios de las tripulaciones, con la finalidad de respetar sus reglas de trabajo.

ScheduleMiser es el motor subyacente que impulsa al sistema de planeación más grande conocido como Flight Ops. ScheduleMiser se basa en una formulación matemática de enteros del problema mezclados. La función objetivo consiste en cinco términos que delinean los diferentes costos del sistema. Se incluyen varios conjuntos de restricciones para asegurar que se satisfagan las demandas, que se programe apropiadamente a las tripulaciones y que no se saturen los aeroplanos. Este sistema fue adoptado e implementado por Raytheon Travel Air en noviembre de 2000 (actualmente Flight Options) para programar su flota de más de 100 aeronaves. Raytheon reportó un ahorro de más de $4.4 millones en el primer año de la implementación de tal sistema. Éste es solamente un ejemplo de muchos sistemas de programación con base matemática que se han implementado en la industria de las aerolíneas.

[1] Martin, C., D. Jones, y P. Keskinocak. "Optimizing On-Demand Aircraft Schedules for Fractional Aircraft Operators" en *Interfaces*, 33, no. 5, septiembre – octubre de 2003, pp. 22-35.

Problemas para la sección 8.6

6. Considere la información dada en el problema 4. Determine la secuencia de la descarga de los camiones para minimizar

   a) El tiempo de flujo medio.
   b) El retraso máximo.
   c) El número de trabajos retardados.

7. El 1 de mayo, un perezoso estudiante de maestría repentinamente se da cuenta de que no ha hecho nada de las tareas y proyectos que le han pedido en siete materias, y el plazo de entrega está por vencerse. Él estima el tiempo requerido para terminar cada proyecto (en días) y también anota sus fechas de entrega:

Proyecto	1	2	3	4	5	6	7
Tiempo (días)	4	8	10	4	3	7	14
Tiempo de entrega	4/20	5/17	5/28	5/28	5/12	5/7	5/15

Ya que los proyectos 1, 3 y 5 son de la misma materia, decide hacerlos en la secuencia en que se vencen. Además, el proyecto 7 requiere resultados de los proyectos 2 y 3, de modo que 7 debe hacerse después de terminar 2 y 3. Determine la secuencia en la que debe realizar los proyectos con objeto de minimizar el retardo máximo.

8. Deben procesarse ocho trabajos en una sola máquina. Los tiempos de procesamiento y las fechas de vencimiento se dan aquí.

Trabajo	1	2	3	4	5	6	7	8
Tiempo de procesamiento	2	3	2	1	4	3	2	2
Fecha de vencimiento	5	4	13	6	12	10	15	19

Además, suponga que deben satisfacerse las siguientes relaciones de precedencia:

$$2 \rightarrow 6 \rightarrow 3.$$

$$1 \rightarrow 4 \rightarrow 7 \rightarrow 8.$$

Determine la secuencia en que estos trabajos deben realizarse para minimizar el retardo máximo sujeto a las restricciones de precedencia.

9. Jane Reed hornea panes y pasteles en su casa que vende para fiestas y otras celebraciones. Jane tiene solamente un horno. Un lunes por la mañana se da cuenta de que se comprometió a terminar cinco pedidos para ese día. Su esposo John hace las entregas; y en cada una tarda aproximadamente 15 minutos. Suponga que ella comienza a hornear a las 8:00 a.m.

Trabajo	**Tiempo requerido**	**Hora prometida**
1	1.2 h	11:30 a.m.
2	40 min	10:00 a.m.
3	2.2 h	11:00 a.m.
4	30 min	1:00 p.m.
5	3.1 h	12:00 h
6	25 min	2:00 p.m.

Determine la secuencia en la que ella debe desempeñar los trabajos con objeto de minimizar

a) El tiempo de flujo medio.

b) El número de trabajos retardados.

c) El retraso máximo.

10. En una sola máquina deben realizarse siete trabajos. Los tiempos de procesamiento y las fechas de vencimiento se dan aquí.

Trabajo	1	2	3	4	5	6	7
Tiempo de procesamiento	3	6	8	4	2	1	7
Fecha de vencimiento	4	8	12	15	11	25	21

Determine la secuencia de los trabajos con objeto de minimizar

a) El tiempo de flujo medio.

b) Número de trabajos retardados.

c) Retraso máximo.

d) ¿Cuál es la terminación de cualquier secuencia?

8.7 ALGORITMOS DE SECUENCIA PARA MÁQUINAS MÚLTIPLES

Ahora ampliamos el análisis de la sección 8.6 al caso en el que varios trabajos deben procesarse en más de una máquina. Suponga que n trabajos deben procesarse a través de m máquinas. El número de horarios posibles es sorprendente, incluso para valores moderados tanto de n como de m. Para cada máquina, existen $n!$ diferentes ordenaciones de los trabajos. Si los trabajos pueden procesarse en las máquinas en cualquier orden, se concluye que existe un total de $(n!)^m$ horarios posibles. Por ejemplo, para $n = 5$, $m = 5$, hay $24{,}883 \times 10^{10}$, o aproximadamente 25 billones de horarios posibles. Aun con la disponibilidad informática actual, resulta imposible o en el mejor de los casos impráctico enumerar todos los horarios posibles, hasta para problemas de tamaño moderado.

En esta sección vamos a presentar algunos resultados conocidos para la programación de trabajos en más de una máquina. Una manera conveniente de representar un horario es mediante una gráfica de Gantt. Como ejemplo, suponga que dos trabajos, I y J, deben programarse en dos máquinas, 1 y 2. Los tiempos de procesamiento son

	Máquina 1	Máquina 2
Trabajo I	4	1
Trabajo J	1	4

Suponga que ambos trabajos deben procesarse primero en la máquina 1 y luego en la máquina 2. Los horarios posibles aparecen en cuatro gráficas de Gantt en la figura 8-6. Los dos primeros horarios se conocen como horarios de permutación. Eso implica que los trabajos se procesan en la misma secuencia en ambas máquinas. Es evidente en este ejemplo que los horarios de permutación suministran un mejor desempeño del sistema en términos de tiempos de flujo total y promedio.

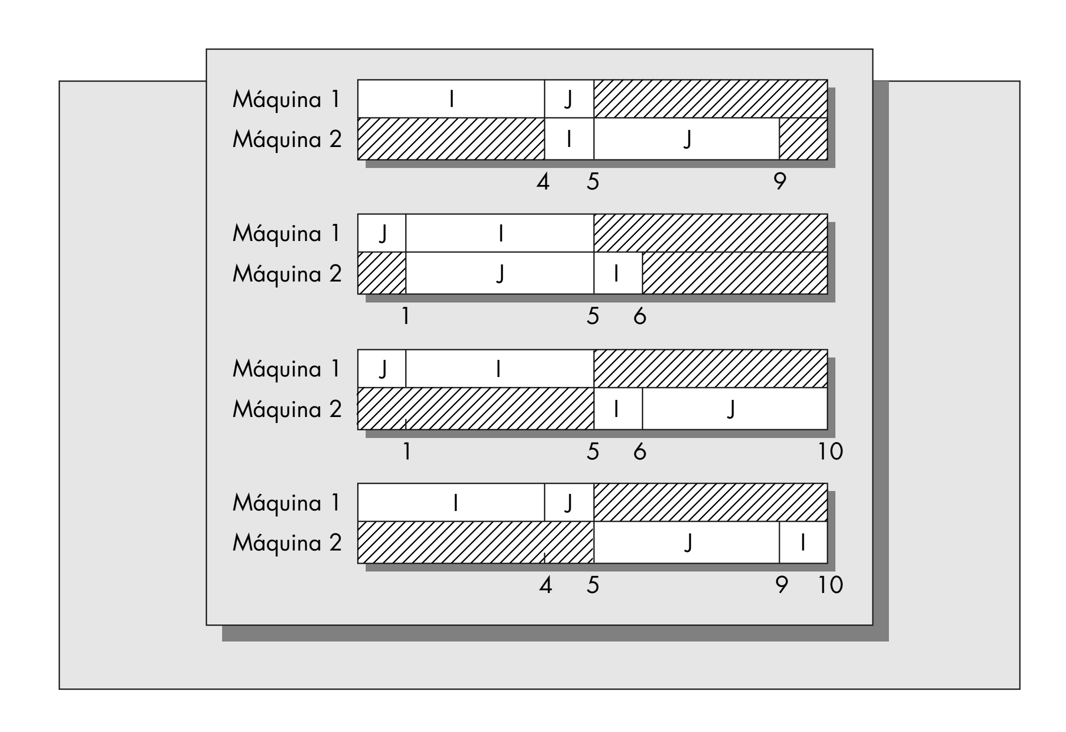


FIGURA 8-6
Todos los horarios posibles para dos trabajos en dos máquinas.

Recuerde que el tiempo de flujo total (o terminación) es el transcurrido desde el inicio del primer trabajo en la primera máquina hasta la terminación del último trabajo en la última máquina. Para los horarios dados, las terminaciones (tiempo de flujo total) son respectivamente 9, 6, 10 y 10.

El tiempo de flujo medio también se usa como una medida del desempeño del sistema. Para el primer horario del ejemplo, el tiempo de flujo medio es $(5 + 9)/2 = 7$. Para el segundo horario, es $(5 + 6)/2 = 5.5$, etcétera.

Un tercer objetivo posible es reducir el tiempo muerto medio del sistema. Éste es el promedio aritmético de los tiempos muertos de cada máquina. En el horario 1, vemos que la máquina 1 está inactiva durante 4 unidades de tiempo (entre los instantes de tiempo 5 y 9) y la máquina 2 también está inactiva durante 4 unidades de tiempo (entre los instantes de tiempo 0 y 4). Entonces, el tiempo muerto medio para el horario 1 es 4. En el horario 2, ambas máquinas 1 y 2 están inactivas durante 1 unidad de tiempo, dando un tiempo muerto medio de 1. Los tiempos muertos medios para los horarios 3 y 4 son 5 unidades de tiempo.

Programación de *n* trabajos en dos máquinas

Suponga que *n* trabajos deben procesarse a través de dos máquinas y que cada trabajo debe procesarse en el orden máquina 1 y luego máquina 2. Además, suponga que el criterio de optimización es minimizar la terminación. El problema de la programación en dos máquinas tiene una solución relativamente simple.

Teorema 8.2 **La solución óptima para programar *n* trabajos en dos máquinas siempre es un horario de permutación.**

El teorema 8.2 implica que puede restringirse la atención a horarios en los cuales la secuencia de trabajos es la misma para ambas máquinas. Es posible demostrar este resultado como se indica en seguida. Considere un horario para *n* trabajos en dos máquinas en las cuales la secuencia de los trabajos es diferente. Es decir, el horario aparece así:

Máquina 1	. . .	I	. . .	J				
Máquina 2					. . .	J	. . .	I

Al invertir la posición de estos trabajos en cada una de las dos máquinas, disminuye el tiempo de flujo. Al programar los trabajos en el orden I-J en la máquina 2 el par (I, J) puede comenzar después de la terminación de I en la máquina 1, en vez de tener que esperar hasta la terminación de J en la máquina 1.

Ya que el número total de horarios de permutación es exactamente $n!$, la determinación de horarios óptimos para dos máquinas tiene aproximadamente el mismo grado de dificultad que la determinación de los horarios óptimos para una máquina.

Un algoritmo muy eficiente para la solución del problema de dos máquinas fue descubierto por Johnson (1954). Éste denota a las maquinas como A y B. Se supone que los trabajos deben procesarse primero en la máquina A y luego en la máquina B. Suponga que los trabajos están rotulados i, para $1 \leq i \leq n$; defina

$$A_i = \text{Tiempo de procesamiento del trabajo } i \text{ en la máquina A.}$$

$$B_i = \text{Tiempo de procesamiento del trabajo } i \text{ en la máquina B.}$$

El resultado de Johnson es que la siguiente regla es óptima para el establecemiento de un orden para procesar los trabajos en las dos máquinas.

Regla: El trabajo i precede al trabajo $i + 1$ si $\text{mín}(A_i, B_{i+1}) < \text{mín}(A_{i+1}, B_i)$.

Una manera fácil de implementar esta regla se indica a continuación:

1. Liste los valores de A_i y B_i en dos columnas.
2. Encuentre el elemento restante más pequeño en las dos columnas. Si aparece en la columna A, entonces programe a ese trabajo como el siguiente. Si aparece en la columna B, entonces programe ese trabajo como el último.
3. Entrecruce los trabajos tal como están programados. Deténgase cuando todos los trabajos hayan sido programados.

Ejemplo *8.5* Cinco trabajos van a programarse en dos máquinas. Los tiempos de procesamiento son

Trabajo	Máquina A	Máquina B
1	5	2
2	1	6
3	9	7
4	3	8
5	10	4

El primer paso consiste en identificar el tiempo mínimo para el trabajo. Es 1, para el trabajo 2 en la máquina A. Ya que aparece en la columna A, el trabajo 2 se programa primero y el renglón 2 se cancela. El siguiente tiempo de procesamiento más pequeño es 2, para el trabajo 1 en la máquina B. Éste aparece en la columna B, de modo que el trabajo 1 se programa como el último. El siguiente tiempo de procesamiento mas pequeño es 3, correspondiente al trabajo 4 en la columna A, de modo que el trabajo 4 se programa como el siguiente. Al continuar de esta manera, obtenemos la secuencia óptima

2–4–3–5–1.

El diagrama de Gantt para el horario óptimo se ilustra en la figura 8-7. Observe que no hay tiempo muerto entre los trabajos de la máquina A. Ésta es una característica de todos los horarios óptimos.

Ampliación a tres máquinas

El problema de programar trabajos es considerablemente más complejo. Si restringimos nuestra atención solamente al tiempo de flujo total, todavía es verdad que un horario de permutación es óptimo (éste no es necesariamente el caso para el tiempo de flujo promedio). Rotule las máquinas como A, B, y C. El problema de las tres máquinas puede reducirse a (esencialmente) un problema de dos máquinas si se satisface la siguiente condición:

$$\text{mín}\, A_i \geq \text{máx}\, B_i \qquad \text{o} \qquad \text{mín}\, C_i \geq \text{máx}\, B_i$$

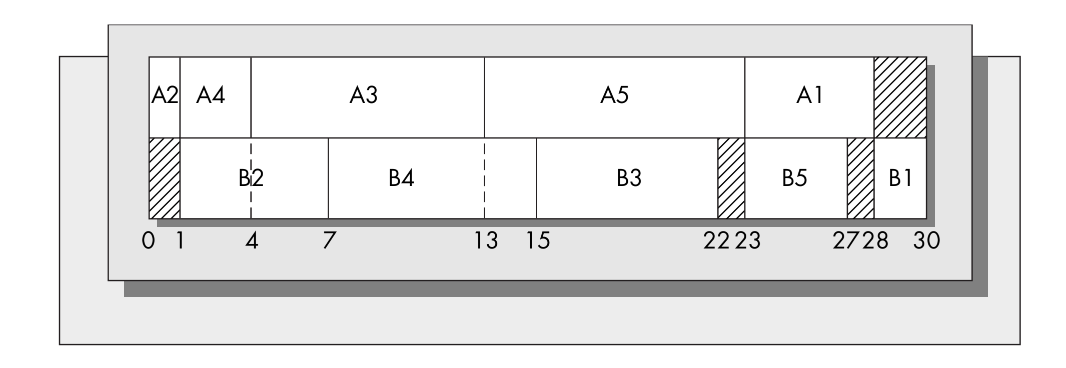


FIGURA 8-7
Diagrama de Gantt para el horario óptimo del ejemplo 8.5.

Solamente es necesario que se cumpla *alguna de estas dos* condiciones. Si ese es el caso, entonces el problema se reduce a un problema de dos máquinas de la siguiente manera.

Defina $A_i' = A_i + B_i$ y $B_i' = B_i + C_i$. Ahora resuelva el problema usando las reglas descritas para dos máquinas, tratando a A_i' y B_i' como los tiempos de procesamiento. El horario de permutación resultante será el óptimo para el problema de las tres máquinas.

Ejemplo *8.6*

Considere los siguientes tiempos de trabajo para un problema de tres máquinas. Suponga que los trabajos se procesan en la secuencia A-B-C.

	Máquina		
Trabajo	**A**	**B**	**C**
1	4	5	8
2	9	6	10
3	8	2	6
4	6	3	7
5	5	4	11

Verificando las condiciones, encontramos

$$\text{mín } A_i = 4,$$
$$\text{máx } B_i = 6,$$
$$\text{mín } C_i = 6,$$

de modo que se satisfaga la condición requerida. Ahora formamos las dos columnas A′ y B′.

	Machine	
Trabajo	**A′**	**B′**
1	9	13
2	15	16
3	10	8
4	9	10
5	9	15

Ahora el problema se resuelve usando el algoritmo de las dos máquinas. La solución óptima es

1–4–5–2–3.

Observe que debido a los empates en la columna A, la solución óptima no es única.

Si no se satisfacen las condiciones para reducir un problema de tres máquinas a un problema de dos máquinas, generalmente este método va a dar resultados razonables, pero posiblemente por debajo de lo óptimo. Siempre que el objetivo sea minimizar la terminación o tiempo de flujo total, un horario de permutación es óptimo para programar tres máquinas. (Sin embargo, no es necesariamente cierto que un horario de permutación sea el óptimo para tres máquinas cuando se use el criterio del tiempo de flujo promedio.)

Observe que suponemos que las máquinas son diferentes y que el procesamiento se desarrolla en forma secuencial: se supone que todos los trabajos se procesan primero en la máquina 1, y después en la máquina 2. Por ejemplo, la máquina 1 puede ser una prensa taladradora y la máquina 2 un torno. Un problema relacionado que estudiamos en el contexto de la programación estocástica es el procesamiento paralelo en máquinas idénticas. En este caso se supone que las máquinas realizan la misma función, y que cualquier trabajo puede ser asignado a cualquier máquina. Por ejemplo, un grupo de 10 trabajos puede requerir procesamiento en alguna de las dos prensas taladradoras. Los resultados del procesamiento paralelo sugieren que

SPT sería una regla efectiva para minimizar el tiempo de flujo medio, pero el tiempo de procesamiento más largo (LPT) con frecuencia tiene mayor efectividad para minimizar el tiempo de flujo total o la terminación. Vamos a estudiar el procesamiento paralelo en el contexto de los tiempos de trabajo aleatorios en la sección 8.8.

El problema del taller de flujo de dos trabajos

Suponga que dos trabajos van a procesarse a través de *m* máquinas. Cada trabajo debe ser procesado por las máquinas en un orden específico, pero las secuencias de los dos trabajos no necesariamente son las mismas. Presentamos un procedimiento gráfico para resolver este problema desarrollado por Akers (1956).

1. Dibuje un sistema de coordenadas cartesianas con los tiempos de procesamiento que corresponden al primer trabajo en el eje horizontal y los tiempos de procesamiento que corresponden al segundo trabajo en el eje vertical. En cada eje, marque los tiempos de operación en el orden de ejecución de las operaciones para ese trabajo.
2. Forme bloques con las áreas que correspondan a cada máquina en la intersección de los intervalos marcados para esa máquina en los dos ejes.
3. Determine una trayectoria desde el origen del extremo del bloque final que no intercepte a ninguno de los bloques y que minimice al movimiento vertical. El movimiento está permitido solamente en tres direcciones: horizontal, vertical y una diagonal a 45 grados. La trayectoria con la distancia vertical mínima indicará la solución óptima. Observe que ésta será igual que la trayectoria con la distancia horizontal mínima.

Este procedimiento se ilustra de la mejor manera con un ejemplo.

Ejemplo ***8.7***

Una fábrica regional de manufacturas produce varios productos para el hogar. Uno de ellos es una lámpara de madera para mesa. Antes de empacarse, las lámparas deben lijarse, barnizarse y pulirse. Cada operación requiere una máquina diferente. En este momento hay dos fletes de dos modelos que esperan el procesamiento. Los tiempos que requieren las tres operaciones para cada uno de los dos fletes son

Trabajo 1		Trabajo 2	
Operación	**Tiempo**	**Operación**	**Tiempo**
Lijado (A)	3	A	2
Barnizado (B)	4	B	5
Pulido (C)	5	C	3

El primer paso es formar los bloques para los tiempos de trabajo en cada uno de los ejes. Véase a la figura 8-8 para este paso. Cada horario factible está representado por una recta que une el origen con la punta del bloque C, a condición de que la recta no atraviese ningún bloque. Solamente se permiten tres tipos de movimientos: horizontal, vertical y una diagonal a 45 grados. El movimiento horizontal implica que solamente se está procesando el trabajo 1, el movimiento vertical implica que solamente se está procesando el trabajo 2 y el movimiento diagonal implica que ambos trabajos están siendo procesados. La minimización del tiempo de flujo es lo mismo que la maximización del tiempo de procesamiento en ambos trabajos. Esto equivale a encontrar la trayectoria desde el origen al extremo del bloque C que maximice el movimiento diagonal y por lo tanto minimice ya sea el movimiento horizontal o el vertical.

En la figura 8-8 se representan dos horarios factibles para este problema. El tiempo total requerido para cualquier horario factible puede obtenerse de dos maneras: ya sea el tiempo total representado en el eje horizontal (12 en este caso) más el movimiento vertical total (4 y 3 respectivamente), o el tiempo total en el eje vertical (10 en este caso) más el movimiento horizontal total (6 y 5 respectivamente). El horario 1 tiene un tiempo total de 16, y el horario 2 tiene un tiempo total de 15. El horario 2 resulta ser el óptimo para este problema.

En la figura 8-9 se presenta el diagrama de Gantt para el horario óptimo.

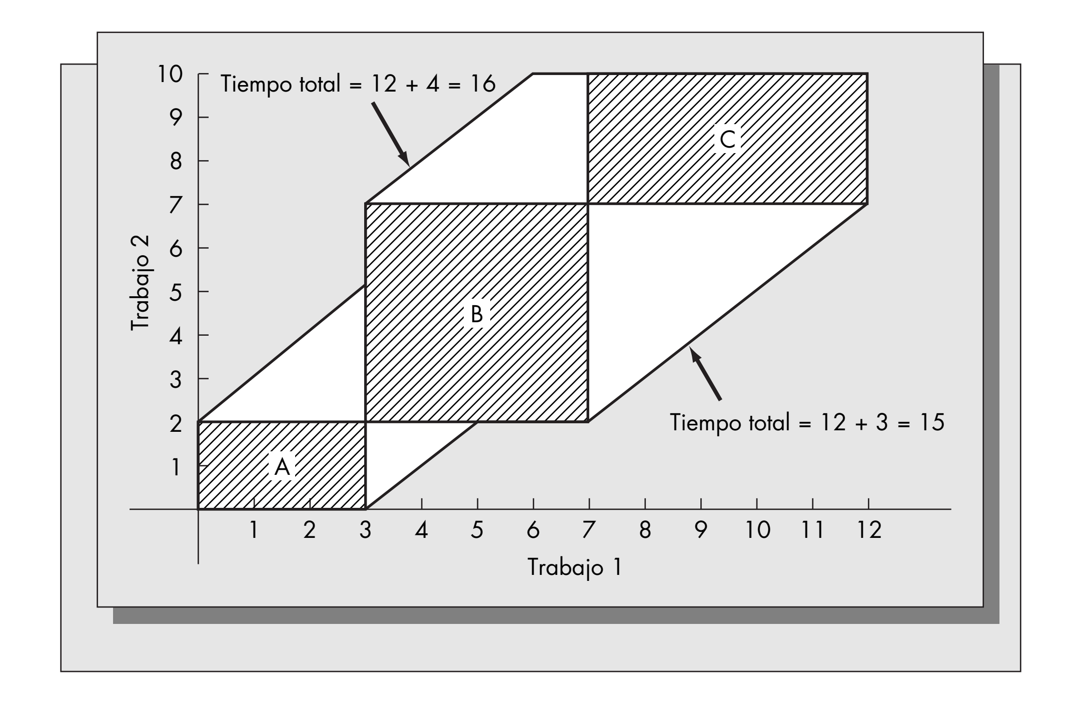


FIGURA 8-8
Solución gráfica del ejemplo 8.7.

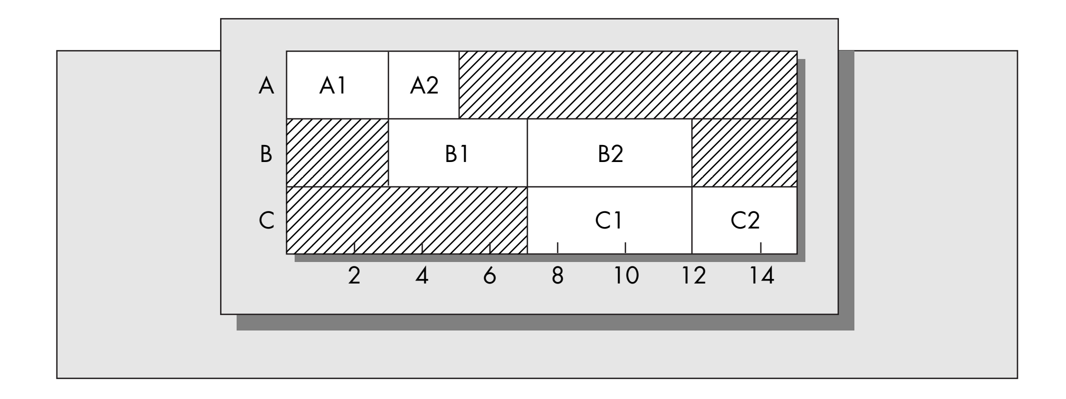


FIGURA 8-9
Diagrama de Gantt para la solución óptima del ejemplo 8.7.

Debemos observar que este método *no* requiere que los dos trabajos se procesen con el mismo orden que el de las máquinas. Presentamos otro ejemplo para ilustrar el caso en el cual difiere la secuencia de los trabajos.

Ejemplo *8.8*

Reggie Sigal y Bob Robinson son compañeros de cuarto; les gusta pasar la mañana del domingo leyendo el periódico. A Reggie le gusta leer primero la sección principal, seguida de la sección deportiva, luego las tiras cómicas, y finalmente los anuncios clasificados. Bob también inicia con la sección principal, pero luego va directamente a los anuncios clasificados, enseguida lee la sección deportiva y finalmente las tiras cómicas. Los tiempos requeridos (en décimas de una hora) para que cada quien lea las diferentes secciones son

Reggie		Bob	
Secuencia requerida	**Tiempo**	**Secuencia requerida**	**Tiempo**
Sección principal (A)	6	Sección principal (A)	4
Deportes (B)	1	Anuncios clasificados (D)	3
Tiras cómicas (C)	5	Deportes (B)	2
Anuncios clasificados (D)	4	Tiras cómicas (C)	5

El objetivo es determinar el orden de lectura de las secciones para que cada quien minimice el tiempo total que se requiere para terminar de leer el periódico. En este problema identificamos a Reggie como el trabajo 1 y a Bob como el trabajo 2. Las secciones del periódico corresponden a las máquinas.

Solución

Con el objeto de obtener la solución óptima, primero formamos los bloques para los tiempos de procesamiento que requiere cada uno de los trabajos. Suponga que se forma el bloque para el trabajo 1 (Reggie) en el eje *x* y el del trabajo 2 (Bob) en el eje *y*. Los tiempos de procesamiento serán secuenciados en cada eje según el orden establecido. La representación gráfica de este problema está dada en la figura 8-10. En dicha figura se representan dos horarios factibles diferentes como trayectorias desde el origen hasta el punto (16, 14). La trayectoria superior representa un horario que requiere que Bob comience a leer primero la sección principal (el trabajo 2 se procesa primero en la máquina A), y la trayectoria inferior requiere que Reggie comience primero. La trayectoria inferior resulta ser la óptima para este problema, con un tiempo de procesamiento de 20.

El tiempo de procesamiento óptimo para este problema es 20; ya que las unidades de tiempo están en décimas de hora, esto es exactamente dos horas. La trayectoria inferior de la figura 8-10 se convierte en un diagrama de Gantt de la siguiente manera: del tiempo 0 al 6, Reggie lee A y Bob está inactivo. Entre los tiempos 6 y 12, la recta de 45 grados indica que tanto Bob como Reggie están leyendo. Reggie lee B durante 1 unidad de tiempo y C durante 5 unidades de tiempo, y Bob lee A durante 4 unidades de tiempo y D durante 3 unidades de tiempo. Ahora Reggie está inactivo durante una unidad de tiempo porque la trayectoria es vertical para este punto, y comienza a leer D en el tiempo 13. Cuando Reggie termina D, él acaba. Comenzando en el tiempo 15, Bob lee C y termina su lectura en el tiempo 20. La figura 8-11 muestra el diagrama de Gantt que indica la solución óptima.

FIGURA 8-10
Solución gráfica del ejemplo 8.8.

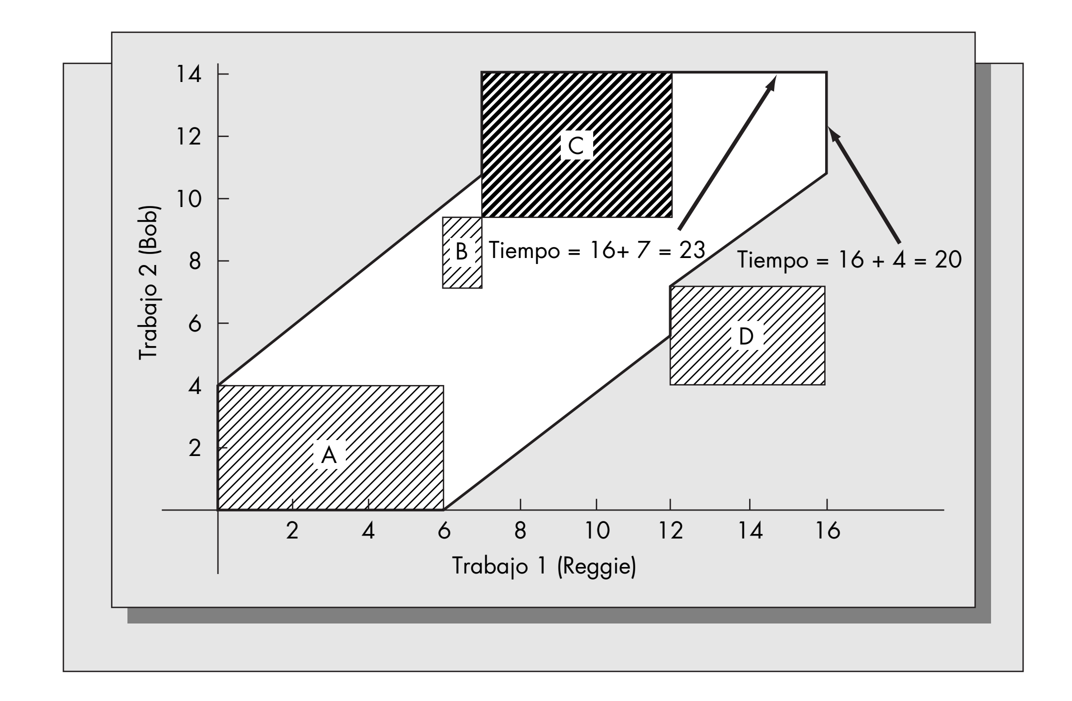
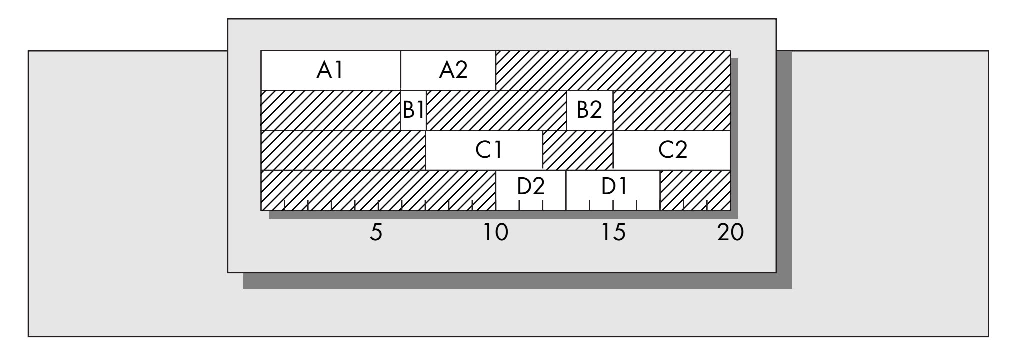


FIGURA 8-11
Diagrama de Gantt para la solución óptima del ejemplo 8.8.



Problemas para la sección 8.7

11. Considere el ejemplo 8.6, que ilustra el uso del algoritmo de Johnson para tres máquinas. Liste todas las soluciones óptimas de este ejemplo.
12. Suponga que deben procesarse 12 trabajos a través de seis máquinas. Si los trabajos pueden procesarse en cualquier orden, ¿cuántos horarios posibles diferentes hay? Si usted fuera a correr un programa de computadora que pudiera evaluar 100 horarios por segundo, ¿cuánto tiempo necesitaría el programa para evaluar todos los horarios factibles?
13. Dos estudiantes de leyes, John y Marsha, están planeando trabajar toda la noche para preparar su examen de leyes al día siguiente. Entre los dos tienen un conjunto de materiales sobre los cinco temas siguientes: contratos, juicios civiles por daños y perjuicios que no se originan en un incumplimiento de contrato, ley civil, ley corporativa y patentes. Con base en su experiencia anterior, ellos estiman que necesitarán los siguientes lapsos (en horas) para cada conjunto de materiales:

	Contratos	**Juicios civiles por daños**	**Derecho civil**	**Derecho corporativo**	**Patentes**
John	1.2	2.2	0.7	0.5	1.5
Marsha	1.8	0.8	3.1	1.1	2.3

Concuerdan en que Marsha tendrá oportunidad de ver cada conjunto de notas antes que John. Suponga que ellos comienzan a estudiar a las 8:00 p. m. Determine los tiempos exactos para que cada quien inicie y termine el estudio de cada materia con objeto de minimizar el tiempo total requerido para que ambos terminen de estudiar las cinco materias.

14. Los siguientes cuatro trabajos deben procesarse en un taller de flujo con tres máquinas.

	Máquina		
Trabajo	**A**	**B**	**C**
1	4	2	6
2	2	3	7
3	6	5	6
4	3	4	8

Encuentre la secuencia óptima de los trabajos con objeto de minimizar el tiempo total de ejecución. ¿Cuál es el tiempo total de ejecución para la solución óptima? Dibuje un diagrama de Gantt que ilustre su solución.

15. Mary y Marcia Brown son dos hermanas que actualmente acuden juntas a la universidad. Cada una requiere tutoría en cinco materias: historia, inglés, matemáticas, ciencias y religión. Ellas estiman que el tiempo (en minutos) que cada una requiere en la tutoría es

	Mary	**Marcia**
Matemáticas	40	20
Historia	15	30
Inglés	25	10
Ciencias	15	35
Religión	20	25

Ellas piensan que los cinco tutores van a estar disponibles todo el día. A Mary le gustaría visitar a los tutores en el orden dado en la tabla, y Marcia preferiría verlos en el orden

siguiente: matemáticas, religión, inglés, ciencias e historia. ¿En qué horario debe planear cada una de ellas ver a los tutores con objeto de minimizar el tiempo total para que ambas terminen sus tutorías?

16. Dos trabajos deben procesarse a través de cuatro máquinas en el mismo orden. Los tiempos de procesamiento en la secuencia requerida son

Trabajo 1		Trabajo 2	
Máquina	**Tiempo**	**Máquina**	**Tiempo**
A	5	A	2
B	4	B	4
C	6	C	3
D	3	D	5

Determine cómo deben programarse los dos trabajos con objeto de minimizar el tiempo total de ejecución y dibuje el diagrama de Gantt que indique el horario óptimo.

17. Peter Minn está planeando ir al Departamento de Tránsito para renovar su licencia de conductor. Su amiga, Patricia, quien lo acompaña, quiere solicitar una nueva licencia. En ambos casos hay cinco pasos requeridos: (A) tomarse una fotografía, (B) firmar un formato de verificación de firma, (C) pasar un examen escrito, (D) pasar un examen de la vista, y (E) pasar un examen de manejo.

    Para las renovaciones, los pasos se ejecutan en el orden A, B, C, D y E con tiempos promedio requeridos, respectivamente, de 0.2, 0.1, 0.3, 0.2 y 0.6 horas. En el caso de solicitudes nuevas, los pasos se ejecutan en la secuencia D, B, C, E y A, con tiempos promedio requeridos de 0.3, 0.2, 0.7, 1.1 y 0.2 horas, respectivamente. Peter y Pat van en un día en que el departamento está más bien vacío. ¿Cómo deben planear su horario con objeto de minimizar el tiempo requerido para que ambos terminen los cinco pasos?

8.8 PROGRAMACIÓN ESTOCÁSTICA: ANÁLISIS ESTÁTICO

Máquina única

Un aspecto que todavía no tratamos es la incertidumbre de los tiempos de procesamiento. En la práctica es posible y aún probable que no pueda predecirse el tiempo exacto de terminación de uno o más trabajos. Es interesante saber si hay o no algunos resultados con respecto a las reglas de secuencia óptima si los tiempos de procesamiento son inciertos. Suponemos que éstos son independientes entre sí.

En el caso del procesamiento en una sola máquina, la mayoría de los resultados son muy similares a los estudiados anteriormente para el caso determinista. Suponga que n trabajos deben procesarse a través de una sola máquina. Suponga que los tiempos de trabajo, $t_1, t_2, \ldots, t_n$, son variables aleatorias con funciones de distribución conocida. La meta es minimizar el tiempo de flujo ponderado promedio *esperado*; es decir,

$$\text{Minimice } E\left(\frac{1}{-}\sum_{i=1}^{n} u_i F_i\right),$$

donde u_i son los pesos y F_i es el tiempo de flujo (promedio) del trabajo i.

Rothkopf (1966) ha demostrado que la solución óptima consiste en ordenar los trabajos de modo que el trabajo i preceda al trabajo $i + 1$ si

$$\frac{E(t_i)}{} < \frac{E(t_{i+1})}{u_{i+1}}.$$

Observe que si establecemos todos los pesos como $u_i = 1$, entonces esta regla consiste simplemente en ordenar los trabajos según el tiempo de procesamiento esperado mínimo; es decir, equivale a SPT.

En el caso de la programación de fechas de entrega con tiempos de procesamiento aleatorios, los resultados también son similares al caso determinista. Banerjee (1965) demuestra que si el objetivo es minimizar al máximo la probabilidad de que un trabajo se atrase, entonces el horario óptimo sería ordenar los trabajos de acuerdo con la fecha de compromiso más próxima (o de acuerdo a la fecha de compromiso esperada más próxima cuando las fechas de compromiso son aleatorias).

Máquinas múltiples

Existen resultados un poco más interesantes para la programación de trabajos con tiempos de trabajo aleatorios en máquinas múltiples. De hecho, existen resultados disponibles para este caso que *no* son verdaderos para el problema determinista.

Una hipótesis que generalmente se formula para el problema de las máquinas múltiples reside en que la distribución de los tiempos de trabajo es exponencial. Esta hipótesis impera porque la distribución exponencial es la única que tiene la *propiedad sin memoria*. (Esta propiedad se estudia con detalle en el capítulo 12 sobre confiabilidad y mantenimiento.) El requerimiento de que los tiempos de trabajo tengan distribución exponencial es riguroso en el contexto de la programación. En la mayoría de los talleres de trabajos, si los tiempos de procesamiento no pueden pronosticarse con exactitud, es poco probable que tengan distribución exponencial. ¿Por qué? La propiedad sin memoria nos dice que la probabilidad de que un trabajo se termine en el siguiente instante de tiempo es independiente del lapso que ya transcurrió al procesar el trabajo. Ciertas aplicaciones, tales como los sistemas telefónicos y las aplicaciones compartidas de computación, pueden modelarse con exactitud de esta manera, pero de ninguna manera la ley exponencial describiría con exactitud a los tiempos de procesamiento para la mayoría de los talleres de trabajos de manufacturas.

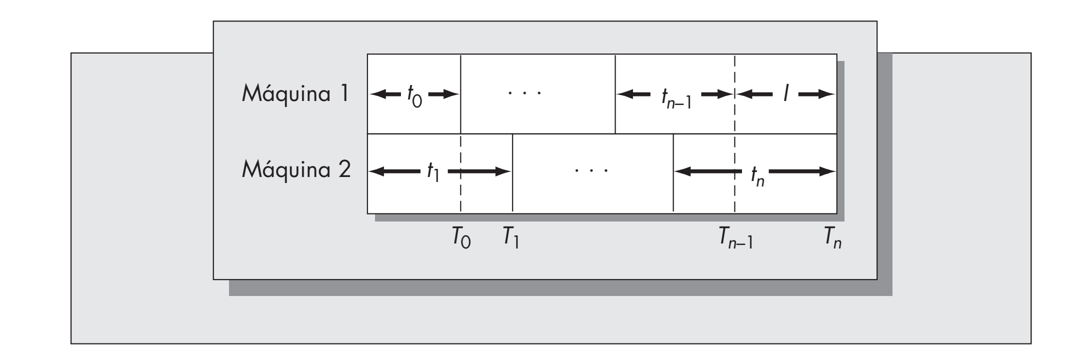
Con estas advertencias, presentamos varios resultados para programar trabajos en máquinas múltiples con tiempos de trabajo aleatorios. Considere el siguiente problema: n trabajos deben procesarse a través de dos máquinas paralelas idénticas. Cada trabajo necesita ser procesado solamente una vez en cada una de las dos máquinas. El objetivo es minimizar el tiempo esperado que transcurre desde el tiempo cero hasta que el último trabajo haya terminado su procesamiento. Esto se conoce como la terminación esperada. Suponemos que los n trabajos tienen tiempos de procesamiento $t_1, t_2, \ldots, t_n$, que son variables aleatorias exponenciales con tasas $\mu_1, \mu_2, \ldots, \mu_n$. Esto implica que el tiempo requerido para terminar el trabajo i, $E(t_i)$, es $1/\mu_i$.

El procesamiento paralelo es diferente del procesamiento del taller de flujo. En el procesamiento del taller de flujo, los trabajos se procesan primero en la máquina 1 y luego en la máquina 2. En el procesamiento paralelo se necesita procesar los trabajos solamente en una máquina, y cualquier trabajo puede procesarse en cualquiera de las dos máquinas. Suponga que para el instante $t = 0$ la máquina 1 está ocupada con un trabajo anterior, el trabajo 0, y el tiempo de procesamiento restante del trabajo 0 es t_0, que podría ser una variable determinista o aleatoria. Los trabajos restantes se procesan como sigue: sea [1], [2], . . . , [n] una permutación de los n trabajos. Se programa el trabajo [1] para la máquina vacante. El trabajo [2] sigue ya sea al trabajo 0 en la máquina 1, o bien al trabajo [1] en la máquina 2, dependiendo de cual se termine primero. Entonces cada trabajo sucesivo se programa para la siguiente máquina disponible.

Sean $T_0 \leq T_1 \leq \cdots \leq T_n$ los tiempos de terminación de los trabajos sucesivos. La terminación es el tiempo de conclusión del último trabajo, que se representa como T_n. El valor esperado de la terminación se minimiza usando la regla del primer tiempo de procesamiento esperado más largo (LEPT)[1]. Observe que esto se opone a la regla SPT para una sola máquina. La optimalidad de programar los trabajos en orden decreciente de su tamaño esperado en vez de hacerlo en orden creciente (como indica la regla SPT) quizás sea más un resultado del procesamiento paralelo que de la aleatoriedad de los tiempos de trabajo.

[1] Sin embargo, si consideramos el tiempo de flujo, entonces el tiempo de procesamiento esperado más corto primero (SEPT, *shortest expected processing time first*) minimiza el tiempo de flujo esperado para dos máquinas.

FIGURA 8-12
Ejecución del procesamiento paralelo en dos máquinas con tiempos de trabajo aleatorios.



Mostramos intuitivamente la optimalidad de LEPT para este problema como sigue: considere el diagrama esquemático de la figura 8-12, que proporciona una ejecución específica de los tiempos de procesamiento para una secuencia arbitraria de los trabajos. En la figura 8-12 la variable aleatoria I corresponde al tiempo muerto de la máquina que no procesa el último trabajo. En forma intuitiva, nos gustaría hacer a I tan pequeño como sea posible con objeto de minimizar la terminación esperada. Esto puede mostrarse más rigurosamente como sigue.

De la ilustración, es claro que

$$T_n + T_{n-1} = \sum_{i=0}^{n} t_i$$

y

$$T_n = T_{n-1} - I.$$

Al despejar T_{n-1} de la segunda ecuación y al sustituir en la primera nos da

$$T_n + T_n - I = \sum_{i=0}^{n} t_i$$

o

$$2T_n = \sum_{i=0}^{n} t_i + I.$$

Si Σt_i se fija independientemente de la secuencia de procesamiento, se concluye que minimizar $E(T_n)$ equivale a minimizar $E(I)$. Ya que I se minimiza al reducir el tiempo de procesamiento del último trabajo, programamos los trabajos en orden decreciente con respecto al tiempo de procesamiento esperado. Observe que este resultado *no* necesariamente se transpone al caso del procesamiento paralelo de dos máquinas con tiempos de procesamiento deterministas. Sin embargo, en el caso determinista, la programación en primer lugar del trabajo de mayor duración generalmente dará buenos resultados cuando se minimice el tiempo de ejecución total. SPT es superior para minimizar el tiempo de flujo medio.

Una clase de problemas que no vamos a estudiar, pero para la que existen varios resultados interesantes, son aquéllos donde los trabajos se procesan a través de m procesadores no idénticos y el tiempo de procesamiento no depende del trabajo. Véase Righter (1988) para la caracterización de la regla óptima de programación para este problema.

El caso del taller de flujo de dos máquinas

Una pregunta interesante es si existe una analogía estocástica con el algoritmo de Johnson para programar n trabajos en dos máquinas en una instalación de taller de fluido; es decir, donde cada trabajo debe procesarse primero a través de la máquina 1, luego a través de la máquina 2. El algoritmo de Johnson nos dice que el trabajo i precede al trabajo $i + 1$ si

$$\text{mín}(A_i, B_{i+1}) < \text{mín}(A_{i+1}, B_i)$$

con objeto de minimizar la terminación.

Ahora suponga que $A_1, A_2, \ldots, A_n$ y $B_1, B_2, \ldots, B_n$ son variables aleatorias exponenciales con tasas respectivas $a_1, a_2, \ldots, a_n$ y $b_1, b_2, \ldots, b_n$. Ahora deseamos minimizar el valor esperado de la terminación. Como el mínimo de dos variables aleatorias exponenciales tiene una tasa igual a la suma de las tasas, se concluye que

$$E[\text{mín}(A_i, B_{i+1})] = \frac{1}{a_i + b_{i+1}}.$$
$$E[\text{mín}(A_{i+1}, B_i)] = \frac{1}{a_{i+1} + b_i}.$$

Por consiguiente, la condición de Johnson se traduce en el caso estocástico a condición de que

$$a_i - b_i \geq a_{i+1} - b_{i+1},$$

de modo que los trabajos deben programarse en el orden de valores decrecientes de la diferencia de las tasas.

Ejemplo *8.9*

Considere el ejemplo 8.5, que se usa para ilustrar el algoritmo de Johnson, pero supongamos que los tiempos de los trabajos son variables aleatorias que tienen la distribución exponencial con tiempos medios dados en el ejemplo. Así tenemos lo siguiente:

	Tiempos de espera		Tasas		
Trabajo	**A**	**B**	**A**	**B**	**Diferencias**
1	5	2	0.20	0.500	−0.30
2	1	6	1.00	0.170	0.83
3	9	7	0.11	0.140	−0.03
4	3	8	0.33	0.125	0.21
5	10	4	0.10	0.250	−0.15

La ordenación de los trabajos de acuerdo con los valores decrecientes de las diferencias de la columna final conduce a la secuencia

2–4–3–5–1.

que es exactamente la misma secuencia que encontramos en el caso determinista usando el algoritmo de Johnson.

Esta sección consideró varios procedimientos de solución cuando los tiempos de trabajo son variables aleatorias. Aun si los tiempos de trabajo se conocen con certeza, todavía puede estar presente la aleatoriedad proveniente de otras fuentes. Por ejemplo, si se considera a la programación como un problema dinámico, debe determinarse el patrón de llegadas al sistema. Es común que los trabajos lleguen de acuerdo con algún proceso aleatorio y que hagan cola para el servicio. La teoría de colas y la simulación son herramientas útiles para manejar la aleatoriedad de este tipo. Conway, Maxwell y Miller (1967) discuten la aplicación de la simulación y de las colas a los problemas de programación de operaciones.

Problemas para la sección 8.8

18. Considere el ejemplo 8.2 de la sección 8.5 para determinar la secuencia óptima de aterrizaje de los aeroplanos. Suponga que los tiempos de aterrizaje son variables aleatorias con una desviación estándar igual a un tercio de la media en cada caso.

a) ¿Cuál debe ser la secuencia de aterrizaje de los aeroplanos con objeto de minimizar el tiempo de flujo ponderado promedio que se espera, si los pesos que van a usarse son los recíprocos del número de pasajeros de cada aeroplano?

b) Para la secuencia que usted encontró en la parte *a*, ¿Cuál es la probabilidad de que todos los aeroplanos aterricen en menos de 100 minutos? Suponga que los tiempos de aterrizaje son variables aleatorias independientes con distribución normal. ¿Cambiará su respuesta si los aeroplanos aterrizan en una secuencia diferente?

19. Un centro de computación tiene dos computadoras idénticas para el proceso por lotes. Éstas se usan como procesadores paralelos. El usuario estima los tiempos de trabajo, pero la experiencia ha mostrado que una distribución exponencial brinda una descripción exacta de los tiempos de trabajo reales. Suponga que para un instante de tiempo hay ocho trabajos pendientes de procesarse con los siguientes tiempos de trabajo esperados (expresados en minutos):

Trabajo	1	2	3	4	5	6	7	8
Tiempo esperado	4	8	1	50	1	30	20	6

a) ¿Cuál debe ser la secuencia de procesamiento de los trabajos con objeto de minimizar el tiempo de terminación esperado de los ocho trabajos (es decir, la terminación)?

b) Suponga que la computadora A está ocupada con un trabajo que tiene exactamente dos minutos de tiempo de procesamiento restante, y que la computadora B está inactiva. Si los tiempos de trabajo son deterministas, muestre los tiempos de inicio y de terminación de cada trabajo en cada computadora usando la secuencia obtenida en la parte *a*.

20. Seis barcos están atracados en una bahía esperando que los descarguen. Los tiempos requeridos para descargarlos son variables aleatorias con medias respectivas de 0.6, 1.2, 2.5, 3.5, 0.4 y 1.8 horas. A los barcos se les da una ponderación prioritaria basada en el tonelaje. Los tonelajes respectivos son 12, 18, 9, 14, 4 y 10. ¿Cuál debe ser la secuencia en que se descarguen los barcos para minimizar el tiempo ponderado esperado?

21. Resuelva el problema 13 suponiendo que los tiempos requeridos por John y Marsha son variables aleatorias con distribución exponencial de los tiempos esperados en el problema 13.

22. Cinco hermanas de una fraternidad femenina planean asistir a una función social. Cada una requiere un peinado de salón de belleza y probarse un vestido. Suponga que los tiempos requeridos son variables aleatorias con distribución exponencial con tiempos medios para el probador de 0.6, 1.2, 1.5, 0.8 y 1.1 horas, respectivamente, y tiempos medios para el peinado de 0.8, 1.6, 1.0, 0.7 y 1.3 horas, respectivamente. Suponga que las pruebas se hacen antes de los peinados y que se dispone solamente de un estilista y una modista. ¿Con qué secuencia deben programarse si se quiere minimizar el tiempo esperado total que se necesita para las pruebas y los peinados?

8.9 PROGRAMACIÓN ESTOCÁSTICA: ANÁLISIS DINÁMICO

Los algoritmos de programación estudiados hasta ahora en este capítulo se basan en la hipótesis de que todos los trabajos llegan al procesamiento simultáneamente. Sin embargo, en la práctica, la programación de trabajos en las máquinas es un problema dinámico. Usamos aquí el término *dinámico* para implicar que los trabajos llegan aleatoriamente con el tiempo y deben tomarse decisiones sobre la marcha en cuanto a cómo programar estos trabajos.

La teoría de colas suministra un medio para modelar algunos problemas de programación dinámica. El suplemento 2, que sigue a este capítulo, suministra un repaso de la teoría básica de colas. En esta sección se supone que los resultados presentados en el suplemento 2 le son familiares.

Considere el siguiente problema. Los trabajos llegan en forma completamente aleatoria a una sola máquina. Esto implica que el proceso de llegada es un proceso de *Poisson*. Suponga que la tasa media de llegada es λ. Inicialmente vamos a suponer que los tiempos de procesamiento tienen una distribución exponencial con media $1/\mu$. Esto significa que la tasa de procesamiento promedio es μ y que los tiempos de procesamiento son variables aleatorias independientes con distribución exponencial idéntica. Finalmente suponemos que los trabajos se procesan sobre la base de primeras llegadas, primeras salidas (FCFS). En la terminología de colas, estamos suponiendo una cola M/M/1/FCFS. También se considerarán otras secuencias de procesamiento.

La teoría básica de colas responde varias preguntas acerca de las características del desempeño de este problema de programación. Primero, se sabe que la distribución de probabilidad del número de trabajos en el sistema (el número de los que esperan ser procesados más el número de los que están siendo procesados) es geométrica con parámetro $\rho = \lambda/\mu$. Es decir, si L es el número de trabajos en el sistema en estado permanente, entonces

$$P\{L = i\} = \rho^i(1 - \rho) \qquad \text{para } i = 0, 1, 2, 3, \ldots$$

El número esperado de trabajos en el sistema es $\rho/(1 - \rho)$. Esto implica que existe una solución solamente para $\rho < 1$. Intuitivamente esto tiene sentido: la tasa de llegada de los trabajos en el sistema debe ser menor que la tasa de procesamiento para garantizar que la cola no crezca sin límite.[2]

Minimizar el tiempo de flujo medio es un objetivo común no solamente en la programación estática, sino también en la programación dinámica. El tiempo de flujo de un trabajo comienza en el instante en que el trabajo se incorpora a la cola de los trabajos sin procesar y continúa hasta que termina su procesamiento. Para el problema de la programación dinámica, el tiempo de flujo de un trabajo es una variable aleatoria; está sujeto a la ejecución de los tiempos de procesamiento de los trabajos precedentes así como a su propio tiempo de procesamiento. El término cola para el tiempo de flujo de un trabajo es el tiempo de espera del sistema y se denota por el símbolo W. El suplemento 2 muestra que la distribución del tiempo de flujo para la cola M/M/1/FCFS es exponencial con parámetro $\mu - \lambda$. Es decir,

$$P\{W > t\} = e^{-(\mu-\lambda)t} \qquad \text{para todo } t > 0.$$

También se obtienen la distribución del tiempo de espera en la cola y el número esperado de trabajos en la cola que aguardan a ser procesados.[3] Podemos ver cómo se aplican estas fórmulas al problema de programación dinámica en el siguiente ejemplo.

Ejemplo *8.10*

Un laboratorio de computación para estudiantes tiene una sola impresora láser. Los trabajos hacen cola para la impresora desde el servidor de la red en el laboratorio y se terminan sobre la base de primeras llegadas, primeras salidas (FCFS). El trabajo de impresión promedio requiere cuatro minutos, pero los tiempos varían considerablemente. La experiencia ha demostrado que la distribución de los tiempos se adhiere a una distribución exponencial. En los tiempos pico, aproximadamente 12 estudiantes por hora requieren el uso de la impresora, pero puede suponerse que la llegada de los trabajos a la impresora ocurre en forma completamente aleatoria.

Suponiendo un periodo de tráfico pico, determine lo siguiente:

a) El número promedio de trabajos en la cola de la impresora.

b) El tiempo de flujo promedio de un trabajo.

---

[2] Sin embargo, lo que no es obvio es lo que sucede en la frontera cuando $\lambda = \mu$ o $\rho = 1$. Ocurre que si las tasas de procesamiento y de llegada son iguales (y los tiempos entre llegadas y los tiempos de procesamiento de las tareas son aleatorios), la cola todavía crece sin límite. Ciertamente la razón de esto no es obvia, pero resulta ser una consecuencia de la aleatoriedad de las llegadas y de los tiempos de procesamiento.

[3] Por conveniencia empleamos el mismo símbolo para una variable aleatoria y su media, usando W tanto para el tiempo de espera en el sistema como para el tiempo de espera previsto en el sistema. Lo mismo se aplicará a L. La razón queda clara a partir del contexto.

c) La probabilidad de que un trabajo espere más de 30 minutos antes de que comience a ser procesado.
d) La probabilidad de que haya más de 6 trabajos en el sistema.

Solución

Primero debemos determinar las tasas de servicio y de llegada. Como cada trabajo de impresión requiere un promedio de 4 minutos, se concluye que la tasa de servicio es $\mu = 1/4$ por minuto o sea 15 por hora. La tasa de llegada está dada como $\lambda = 12$ por hora. La intensidad de tráfico es $\rho = 12/15 = .8$.

a) El número promedio de trabajos en la cola es L_q, que está dado por (consúltese el suplemento 2)

$$L_q = \frac{\rho^2}{1-\rho} = \frac{.64}{.2} = 3.2.$$

b) El tiempo de flujo promedio de un trabajo es el mismo que el tiempo de espera en el sistema. Del suplemento 2,

$$W = \frac{\rho}{\lambda(1-\rho)} = \frac{.8}{(12)(.2)} = 0.3333 \text{ horas (20 minutos).}$$

c) Aquí estamos interesados en $P\{W_q > 0.5\}$. La distribución de W_q es esencialmente exponencial con parámetro $\mu - \lambda$, pero con masa positiva en cero.

$$P\{W_q > t\} = \rho e^{-(\mu-\lambda)t} = .8e^{-(3)(0.5)} = .1785.$$

d) Aquí deseamos determinar $P\{L > 6\}$. De la solución del problema S2.1 en el suplemento 2, demostramos que

$$P\{L > k\} = \rho^{k+1} = .8^7 = .2097.$$

Disciplinas de selección independiente de los tiempos de procesamiento del trabajo

Aun cuando nuestro sentido del juego limpio dice que la disciplina de servicio debe ser FCFS, hay muchas ocasiones en que los trabajos se procesan con otras secuencias. Por ejemplo, considere un proceso de manufactura en el cual las partes se apilan a medida que se terminan. La siguiente etapa del proceso podría ser simplemente tomar las partes de la sección superior de la pila, lo que conduce a una disciplina último en llegar, primero en salir (LCPS: *last-come, first-served*). En forma similar, cabe visualizar una situación en la cual las partes terminadas se arrojan a una tolva y se extraen aleatoriamente, lo que conduce a que el servicio se presente en orden aleatorio.

La teoría de colas nos dice que siempre que la disciplina de selección *no dependa de los tiempos de procesamiento*, los tiempos de flujo medios (y por tanto los números medios en el sistema y las longitudes medias de las colas) son los mismos.[4] Sin embargo, la **varianza** de los tiempos de flujo ciertamente dependerá de la disciplina de selección. La varianza del tiempo de flujo es máxima con LCFS y mínima con FCFS entre las tres disciplinas FCFS, LCFS, y aleatoria.[5] Los segundos momentos del tiempo de flujo están dados por

$$E_{\text{LCFS}}(W^2) = \frac{1}{1-\rho} E_{\text{FCFS}}(W^2),$$

$$E_{\text{ALEATORIA}}(W^2) = \frac{1}{1-\rho/2} E_{\text{FCFS}}(W^2).$$

(La segunda ecuación ha sido probada solamente para tiempos de procesamiento exponencial, pero se ha conjeturado que en general es válida.) Recuerde que la varianza de una

[4] Muchos de los resultados citados en esta sección se discuten con más detalle en el capítulo 8 de Conway, Maxwell y Miller (1967). De hecho, aun cuando este libro tiene más de 30 años de antigüedad, todavía ofrece el tratamiento más comprensivo de la aplicación de la teoría de colas a los problemas de programación.

[5] Esto es estrictamente cierto si eliminamos la posibilidad de la priorización. Priorización significa que a un trabajo que acaba de llegar se le permite interrumpir el servicio de un trabajo que ya está en progreso. No vamos a tratar aquí a las disciplinas priorizantes.

variable aleatoria es el segundo momento menos la media al cuadrado, de modo que podemos obtener la varianza de W directamente de estas fórmulas.

Ejemplo *8.10* *(continuación)*

Determine la varianza de los tiempos de flujo para la impresora láser durante las horas pico suponiendo (1) FCFS, (2) LCFS y (3) disciplinas de selección aleatoria.

Solución

Debido a que bajo FCFS el tiempo de flujo tiene distribución exponencial con parámetro $\mu - \lambda$, se concluye que el tiempo de flujo medio es $1/(\mu - \lambda)$ y la varianza del tiempo de flujo es $1/(\mu - \lambda)^2$.

$$E_{\text{FCFS}}(W) = \frac{1}{\mu - \lambda} = \frac{1}{15 - 12} = \frac{1}{3} \text{ hora.}$$

$$\text{Var}_{\text{FCFS}}(W) = \frac{1}{(3)^2} = \frac{1}{9}.$$

Ya que $\text{Var}(W) = E(W^2) - (E(W))^2$, se concluye que

$$E_{\text{FCFS}}(W^2) = \text{Var}_{\text{FCFS}}(W) + (E_{\text{FCFS}}(W))^2 = \frac{1}{9} + \left(\frac{1}{3}\right)^2 = \frac{2}{9}.$$

A partir de estos resultados, obtenemos

$$E_{\text{LCFS}}(W^2) = \left(\frac{1}{1 - .8}\right)\left(\frac{2}{9}\right) = \frac{10}{9}, \qquad \text{dando}$$

$$\text{Var}_{\text{LCFS}}(W) = \frac{10}{9} - \left(\frac{1}{3}\right)^2 = 1.0.$$

En forma similar,

$$E_{\text{ALEATORIA}}(W^2) = \left(\frac{1}{1 - .4}\right)\left(\frac{2}{9}\right) = 0.3704, \qquad \text{dando}$$

$$\text{Var}_{\text{ALEATORIA}}(W) = 0.3704 - \left(\frac{1}{3}\right)^2 = 0.2593.$$

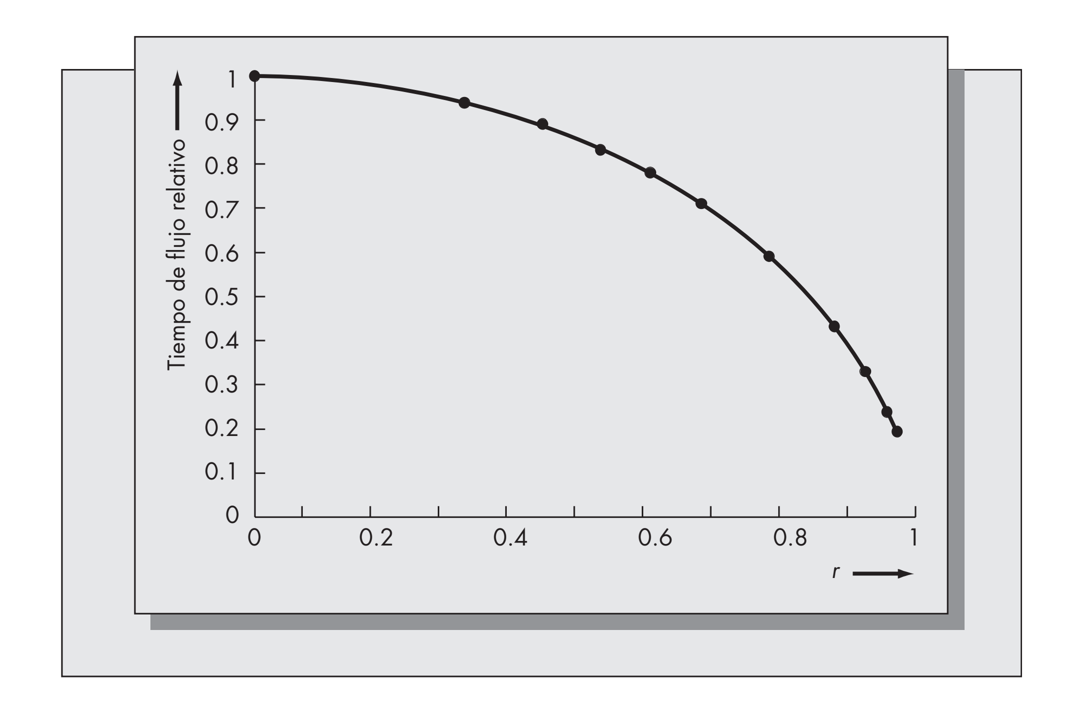
Entonces vemos que, si los trabajos se procesan en la impresora sobre una base LCFS, la varianza del tiempo de flujo es 1.0, en comparación con 2/9 para FCFS. El procesamiento de trabajos en orden aleatorio también incrementa la varianza del tiempo de flujo sobre FCFS, pero en grado mucho menor. Intuitivamente, la varianza es mayor para LCFS que para FCFS, porque cuando los trabajos se procesan en orden opuesto a su llegada, es posible que un trabajo que haya estado en la cola por un tiempo siga siendo "rebotado" por trabajos que acaban de llegar. El resultado será un tiempo de flujo muy largo. Se presenta un fenómeno similar en el caso aleatorio, pero el efecto no es tan severo.

Disciplinas de selección dependientes de los tiempos de procesamiento del trabajo

Uno de los objetivos de la investigación sobre modelos de colas dinámicas reside en descubrir las disciplinas de selección óptima. Anteriormente afirmamos que las medidas promedio del desempeño (W, W_q, L y L_q) son independientes de la disciplina de selección siempre que ésta no se encuentre sujeta a los tiempos de procesamiento de los trabajos. Sin embargo, considere el caso en el que los tiempos de procesamiento de los trabajos se ejecutan en el instante en que el trabajo se incorpora a la cola. Esta hipótesis es razonable para la mayoría de las aplicaciones de programación industrial. Para los problemas de procesamiento de máquinas, es posible que el contenido de trabajo sea un múltiplo del número de partes que deben procesarse, de modo que el tiempo de procesamiento se conoce en el instante en que un trabajo se incorpora a la cola. Un ejemplo donde esto no es válido sería en un banco. Generalmente no es posible decir cuánto

FIGURA 8-13
Tiempos de flujo relativos: SPT contra FCFS.

Fuente: adaptado de los resultados que se tabulan en Conway, Maxwell y Miller (1967), p. 184.



tiempo va a requerir un cliente para realizar sus operaciones hasta que revelen la naturaleza de las mismas.

Si los tiempos de procesamiento de los trabajos se ejecutan cuando un trabajo se incorpora a la cola, es posible usar una disciplina de selección que dependa de los tiempos de los trabajos. Una disciplina de este tipo, que ya se estudió minuciosamente en este capítulo, es la regla SPT: el siguiente trabajo procesado es aquel con el tiempo de procesamiento más corto. Resulta que la regla SPT es efectiva para el problema dinámico así como para el problema estático.

La SPT es capaz de reducir significativamente el tamaño de la cola en el problema de la programación dinámica. En la figura 8-13 mostramos qué tan dramático puede ser dicho efecto. Esta figura supone una cola M/M/1. Defina el tiempo de flujo relativo como la relación

$$\frac{E_{\text{SPT}}(\text{Tiempo de flujo})}{E_{\text{FCFS}}(\text{Tiempo de flujo})}.$$

Vemos que, a medida que aumenta la intensidad del tráfico, mejora la ventaja de SPT para reducir el tiempo de flujo (y por tanto el número en el sistema y la longitud de la cola). La cola se reduce “limpiando” los trabajos más cortos. Para valores de ρ cercanos a 1, esta relación podría ser tan baja como 0.2. Debido a la fórmula de Little (véase el suplemento 2), el número esperado en el sistema y el tiempo de flujo son proporcionales, de modo que esta curva también representa la relación de los números esperados en el sistema para las respectivas disciplinas de selección.

Sería interesante preguntarse qué pasa con la varianza de las medidas de desempeño para cada disciplina de selección. Nuevamente suponiendo tiempos de servicio exponenciales, la tabla 8-1 proporciona las varianzas de los tiempos de flujo para SPT y FCFS como una función de la intensidad del tráfico, ρ.

Lo que vemos de esta tabla es que para valores bajos de la intensidad de tráfico (menor que aproximadamente .7), SPT tiene una varianza ligeramente menor que FCFS. Sin embargo, a medida que la intensidad de tráfico se aproxima a 1, la varianza bajo SPT aumenta dramáticamente. Para $\rho = .99$, la varianza del tiempo de flujo bajo SPT es más de 16 veces el de FCFS. Así, la reducción del tiempo de flujo medio alcanzada por SPT viene al costo de aumentar posiblemente la varianza.

TABLA 8-1 Varianza del tiempo de flujo bajo FCFS y SPT

ρ	FCFS	SPT
.1	1.2345	1.179
.2	1.5625	1.482
.3	2.0408	1.896
.4	2.6666	2.563
.5	4.0000	3.601
.6	6.2500	5.713
.7	11.1111	12.297
.8	25	32.316
.9	100	222.2
.95	400	1 596.5
.98	2 500	22 096
.99	10 000	161 874

Fuente: Conway, Maxwell y Miller (1967), p. 189.

Gran parte de la investigación sobre programación dinámica considera las disciplinas de prioridad, lo que implica que los trabajos entrantes se clasifican en grupos y se da prioridad a ciertos grupos sobre otros. La programación por prioridad es común en las salas de emergencia de los hospitales. También se presenta en la programación de los trabajos por lotes en una computadora *mainframe*, ya que permite dar prioridad a ciertos usuarios sobre otros. Las prioridades pueden ser priorizables o no priorizables. La prioridad priorizable significa que cuando llega un trabajo con una prioridad más alta que el que se está procesando se permite interrumpir el servicio del trabajo en proceso. El trabajo interrumpido puede continuar después donde se detuvo, o tal vez el procesamiento deba iniciar nuevamente a partir de la memoria de "apuntes". No vamos a proseguir con esta área compleja, pero observe que SPT es bastante robusto; es óptimo para una clase grande de problemas de programación de prioridades.

La regla $c\mu$

Considere el siguiente problema de programación. Los trabajos llegan aleatoriamente a una sola máquina con tiempos de procesamiento exponenciales. Permitimos que los trabajos en la cola tengan diferentes tasas de servicio μ_i. Es decir, para cualquier momento, suponga que hay n trabajos que esperan ser procesados. Indexamos los trabajos como 1, 2, 3, . . . , n y suponemos que el tiempo requerido para terminar el trabajo i tiene una distribución exponencial con media $1/\mu_i$. Además, suponga que hay un retorno de c_i si el trabajo i se termina en algún tiempo fijo t. ¿Cuál es la mejor selección para el siguiente trabajo que va a procesarse si el objetivo es maximizar las ganancias totales esperadas?

Derman *et al*. (1978) mostraron que la política óptima consiste en seleccionar el trabajo con el valor máximo de $c_i\mu_i$. Observe que si los pesos son iguales a 1, la regla $c\mu$ tiene exactamente la misma esperanza que la SPT. Así, esto puede considerarse como un tipo de regla de programación ponderada SPT. Resulta que la regla $c\mu$ es óptima para otras versiones del problema de programación estocástica. Remitimos al lector interesado a Pinedo (1983) y a las referencias listadas ahí.

Problemas para la sección 8.9

23. Un sistema de computación forma colas con los trabajos por lotes y los procesa sobre una base FCFS. Entre las 2 y las 5 p.m., los trabajos llegan con una tasa promedio de 30 por hora y en promedio requieren 1.2 minutos de tiempo de cómputo. Suponga que el proceso de llegada es de Poisson y que los tiempos de procesamiento tienen distribución exponencial.

    a) ¿Cuál es el número esperado de trabajos en el sistema y en la cola en el estado permanente?

b) ¿Cuál es el tiempo de flujo esperado y el tiempo en la cola en el estado permanente?
c) ¿Cuál es la probabilidad de que el sistema esté vacío?
d) ¿Cuál es la probabilidad de que la cola esté vacía?
e) ¿Cuál es la probabilidad de que el tiempo de flujo de un trabajo sobrepase los 10 minutos?

24. Considere el sistema de cómputo del problema 23.

a) Calcule la varianza de los tiempos de flujo suponiendo disciplinas FCFS, LCFS y de selección aleatoria.
b) Usando una aproximación normal de la distribución del tiempo de flujo para los casos LCFS y aleatorio, estime la probabilidad de que el tiempo de flujo en el sistema sobrepase los 10 minutos en cada caso.

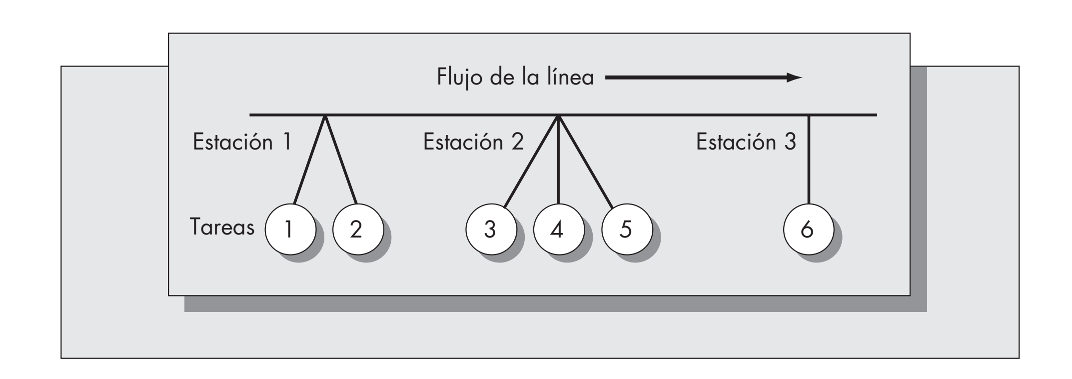
25. Un recurso crítico en una operación de manufactura experimenta una intensidad de tráfico muy alta durante los periodos de mayor ocupación del taller. Durante estos periodos la tasa de llegada es de aproximadamente 57 trabajos por hora. Los tiempos con que se procesan los trabajos tienen una distribución aproximadamente exponencial con media de 1 minuto.

a) Calcule el tiempo de flujo esperado en el sistema suponiendo una disciplina de procesamiento FCFS, y el tiempo de flujo esperado bajo SPT usando la figura 8-13.
b) Calcule la probabilidad de que un trabajo espere por más de 30 minutos para su procesamiento bajo FCFS.
c) Usando una aproximación normal, estime la probabilidad de que un trabajo espere por más de 30 minutos para su procesamiento bajo LCFS.

8.10 BALANCEO DE LÍNEA DE ENSAMBLE

El balanceo de una línea de ensamble es un problema clásico de ingeniería industrial. Aunque gran parte del trabajo en esta área se remonta a mediados de la década de los cincuenta y a inicios de la década de los sesenta, actualmente la estructura básica del problema es relevante para el diseño de los sistemas de producción, incluso en plantas automatizadas. El problema se caracteriza por un conjunto de *n* tareas diferentes que deben terminarse para cada artículo. El tiempo requerido para terminar la tarea *i* es una constante conocida t_i. El objetivo es organizar las tareas en grupos, ejecutándose cada grupo en una sola estación de trabajo. En la mayoría de los casos, la cantidad de tiempo asignada a cada estación de trabajo se determina con antelación, basándose en la tasa deseada de producción de la línea de ensamblaje. Esto se conoce como el tiempo de ciclo y se denota como *C*. En la figura 8-14 se proporciona un esquema de una línea de ensamblaje típica. En la figura, los círculos representan las tareas que van a hacerse en las correspondientes estaciones.

FIGURA 8-14
Esquema de una línea de ensamble típica.



Tradicionalmente se piensa en el balanceo de la línea de ensamble como un problema de disposición y diseño de instalaciones, y podría argumentarse que sería más apropiado incluirlo en el capítulo 10. Tradicionalmente, asignar tareas a las estaciones de trabajo ha sido una decisión de un solo golpe que se toma en el momento de construir la planta e instalar la línea. No obstante, la naturaleza de la fábrica moderna está cambiando. Hay mayor flexibilidad en el diseño de plantas nuevas, y esto permite que surjan nuevas líneas y que las antiguas se reestructuren sobre una base continua. En un ambiente de este tipo, el balanceo de línea se parece más a un problema de programación dinámica que a un problema de disposición de instalaciones en un solo golpe.

Existen varios factores que contribuyen a la dificultad del problema. Primero, hay restricciones de precedencia: algunas tareas deben terminarse según cierta secuencia. Otro problema es que algunas tareas no pueden ejecutarse en la misma estación de trabajo. Por ejemplo, podría no ser posible trabajar en el extremo frontal y en el extremo trasero de un objeto grande tal como un automóvil en la misma estación de trabajo. Esto se conoce como una *restricción de zonificación*. Todavía pueden surgir otras complicaciones. Por ejemplo, ciertas tareas tendrían que terminarse en la misma estación de trabajo, y otras podrían requerir más de un trabajador.

Encontrar el balance óptimo de una línea de ensamble es un problema combinatorio difícil aun cuando no estén presentes los problemas anteriormente descritos. Se han sugerido varias heurísticas relativamente simples para determinar un balance aproximado. Muchos de estos métodos requieren pocos cálculos y hacen posible la solución manual de problemas grandes.

Sea $t_1, t_2, \ldots, t_n$ el tiempo requerido para terminar las respectivas tareas. El contenido total de trabajo asociado con la producción de un artículo, por ejemplo T, está dado por

$$T = \sum_{i=1}^{n} t_i.$$

Para un tiempo de ciclo de C, el número mínimo posible de estaciones de trabajo es $[T/C]$, donde los paréntesis rectangulares indican que el valor de T/C debe redondearse al siguiente entero. Por la naturaleza discreta e indivisible de las tareas y las restricciones de precedencia, frecuentemente es verdad que se requieren más estaciones que este valor mínimo ideal. Si existe un margen para seleccionar el tiempo de ciclo, se aconseja experimentar con diferentes valores de C para ver si puede obtenerse un balance más eficiente.

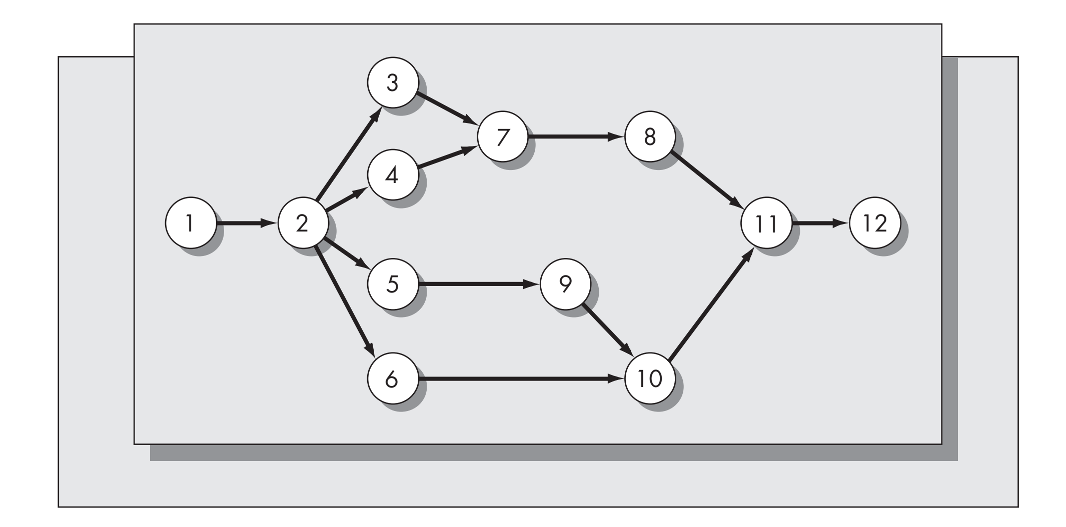
Presentaremos un método heurístico de Helgeson y Birnie (1961) conocido como *la técnica de ordenamiento según el peso posicional*. El método coloca un peso en cada tarea basándose en el tiempo total requerido por todas las tareas subsiguientes. Las tareas se asignan en forma secuencial a las estaciones basadas en estos pesos. Ilustramos el método con un ejemplo.

Ejemplo *8.11*

El ensamble final de las computadoras personales sin marca, un clon genérico de PC que se ordena por correo, requiere un total de 12 tareas. El ensamble se hace en la planta de Lubbock, Texas, usando diferentes componentes importados del extremo Oriente. Las tareas requeridas para las operaciones de ensamble son

1. Taladrar orificios en el gabinete de metal y montar las ménsulas para sostener las unidades de disco.
2. Fijar la tarjeta madre al gabinete.
3. Montar la fuente de poder y unirla con la tarjeta madre.
4. Colocar el procesador principal y los chips de memoria en la tarjeta madre.
5. Enchufar la tarjeta de gráficos.
6. Montar las unidades de discos flexibles. Unir el controlador de la unidad de discos flexibles y la fuente de poder a las unidades de disco.
7. Montar la unidad del disco duro. Unir el controlador del disco duro y la fuente de poder al disco duro.
8. Se hacen las conexiones apropiadas en la tarjeta madre para la configuración específica del sistema.
9. Unir el monitor a la tarjeta de gráficos antes de correr el diagnóstico del sistema.

FIGURA 8-15
Restricciones de precedencia para la computadora sin marca (ejemplo 8.11).



10. Correr el diagnostico del sistema.
11. Sellar el gabinete.
12. Adherir el logo de la compañía y empacar el sistema para su embarque.

Los orificios deben taladrarse y la tarjeta madre conectarse al gabinete antes que cualesquiera otras operaciones. Una vez que se ha montado la tarjeta, pueden instalarse la fuente de poder, la memoria, los chips del procesador, la tarjeta de gráficos y los controladores de discos. Las unidades de discos flexibles se colocan en la unidad antes que la unidad de disco duro, y requieren que se coloque en primer lugar la fuente de poder y el controlador. Basándose en la configuración de la memoria y en la selección del adaptador de gráficos, se determinan y se gradúan los valores de los interruptores en la tarjeta madre. El monitor debe conectarse a la tarjeta de gráficos de modo que puedan leerse los resultados de las pruebas de diagnóstico. Finalmente, después de terminar todas las demás tareas, se corren los diagnósticos y el sistema se empaca para su embarque. En la siguiente tabla se resumen los tiempos de los trabajos y las relaciones de precedencia para este problema. La representación en red de este problema específico se da en la figura 8-15.

Tarea	Predecesores inmediatos	Tiempo
1	—	12
2	1	6
3	2	6
4	2	2
5	2	2
6	2	12
7	3, 4	7
8	7	5
9	5	1
10	9, 6	4
11	8, 10	6
12	11	7

Suponga que la compañía desea contratar suficientes trabajadores para producir una máquina ensamblada cada 15 minutos. La suma de los tiempos de las tareas es 70, lo que implica que el número mínimo de estaciones de trabajo es la relación $70/15 = 4.67$ redondeado al siguiente entero, que es 5. Esto no significa que necesariamente exista un balance entre las cinco estaciones.

El procedimiento de solución requiere que se determine el peso posicional de cada tarea. El peso posicional de la tarea *i* se define como el tiempo requerido para ejecutar la tarea *i* más los tiempos requeridos para ejecutar todas las tareas que tengan a la tarea *i* como su predecesor. Ya que la tarea 1 debe preceder a todas las demás, su peso posicional es simplemente la suma de los tiempos de las tareas, que es 70. La tarea 2 tiene un peso posicional de 58.

TABLA 8-2 Pesos posicionales para el ejemplo 8.11

Tarea	Peso posicional
1	70
2	58
3	31
4	27
5	20
6	29
7	25
8	18
9	18
10	17
11	13
12	7

En la figura 8-15 vemos que la tarea 3 debe preceder a las tareas 7, 8, 11, y 12, de modo que el peso posicional de la tarea 3 es $t_3 + t_7 + t_8 + t_{11} + t_{12} = 31$. Los otros pesos posicionales se calculan de manera similar. Éstos se listan en la tabla 8-2.

El siguiente paso consiste en ordenar las tareas en orden decreciente del peso posicional. La jerarquización para este caso es 1, 2, 3, 6, 4, 7, 5, 8, 9, 10, 11, 12. Finalmente, las tareas se asignan en forma secuencial a las estaciones en el orden de jerarquización y se hacen asignaciones siempre que no se violen las restricciones de precedencia.

Consideremos ahora el balance obtenido con esta técnica suponiendo un tiempo de ciclo de 15 minutos. La tarea 1 se asigna a la estación 1. Eso deja una holgura de tres minutos para esta estación. Sin embargo, ya que la siguiente tarea por asignarse es la tarea 2, con objeto de no violar las restricciones de precedencia, y la suma $t_1 + t_2$ sobrepasa a 15, cerramos la estación 1. Entonces se asignan las tareas 2, 3 y 4 a la estación 2, lo que conduce a un tiempo muerto para esta estación de sólo un minuto. Continuando así, obtenemos el siguiente balance para este problema:

Estación	1	2	3	4	5	6
Tareas	1	2, 3, 4	5, 6, 9	7, 8	10, 11	12
Tiempo muerto	3	1	0	3	5	8

Observe que aunque el número mínimo posible de estaciones para este problema es cinco, la técnica de jerarquización de los pesos posicionales conduce a un balance de seis estaciones. Como el método es solamente una heurística, es posible que exista una solución con cinco estaciones. Sin embargo, en este caso, el balance óptimo requiere seis estaciones cuando $C = 15$ minutos.

El gerente de la compañía que ensambla las computadoras sin marca tiene interés en determinar el tiempo de ciclo mínimo que conduciría a un balance de cinco estaciones. Si incrementamos el tiempo de ciclo de $C = 15$ a $C = 16$, entonces el balance obtenido es

Estación	1	2	3	4	5
Tareas	1	2, 3, 4, 5	6, 9	7, 8, 10	11, 12
Tiempo muerto	4	0	3	0	3

Es evidente que éste es un balance mucho más eficiente: el tiempo muerto total ha sido reducido de 20 minutos por unidad a solamente 10 minutos por unidad. El número de estaciones disminuye en aproximadamente 16 por ciento mientras que el tiempo de ciclo sólo aumenta aproximadamente 7 por ciento. Suponiendo que un día de producción tiene siete horas, un valor de $C = 15$ conduciría a un nivel diario de producción de 28 unidades por operación de ensamblado, y un valor de $C = 16$ minutos conduciría a un nivel diario de producción de 26.25 unidades por operación de ensamblado. La administración tendría que determinar si los ahorros obtenidos con cinco en lugar de seis estaciones justifican la declinación de la tasa de producción de 1.75 unidades por día por operación.

Una elección alternativa es quedarse con las seis estaciones, pero veamos si se puede obtener un balance de seis estaciones con un tiempo de ciclo menor que 15 minutos. Resulta que para los valores del tiempo de ciclo tanto de 14 minutos como de 13 minutos, el método de jerarquización del peso posicional va a dar balances de seis estaciones. La solución $C = 13$ es

Estación	1	2	3	4	5	6
Tareas	1	2, 3	6	4, 5, 7, 9	8, 10	11, 12
Tiempo muerto	1	1	1	1	4	0

Resulta que 13 minutos es el tiempo de ciclo mínimo con seis estaciones. El tiempo muerto total de ocho minutos que resulta del balance anterior es dos minutos menor que el que se alcanza con cinco estaciones cuando $C = 16$. La tasa de producción con seis estaciones y $C = 13$ sería 32.3 unidades por día por operación. El incremento del número de estaciones de cinco a seis conduce a una mejora importante en la tasa de producción.

En esta sección presentamos la heurística del peso posicional jerarquizado para resolver el problema de balanceo de la línea de ensamblado. También existen otros métodos heurísticos. Uno es COMSOAL, una heurística basada en la computación desarrollada por Arcus (1966). El método es eficiente para problemas grandes que incluyen muchas tareas y estaciones de trabajo. Kilbridge y Wester (1961) sugieren un método similar a la técnica de jerarquización del peso posicional.

Hay procedimientos óptimos para resolver el problema de balanceo de línea, pero los cálculos son complejos y consumen tiempo, pues requieren ya sea programación dinámica (Held y otros, 1963) o programación entera (Thangavelu y Shetty, 1971). El interés más reciente en el problema de balanceo de línea se ha centrado en aspectos relacionados con la incertidumbre en los tiempos de ejecución de las tareas individuales. (Véase Hillier y Boling, 1986, y las referencias contenidas ahí.)

Virtualmente todos los procedimientos de balanceo de la línea de ensamble suponen que el objetivo es minimizar el tiempo muerto total en todas las estaciones de trabajo. Sin embargo, como vimos en esta sección, un balance óptimo para un tiempo de ciclo fijo puede no ser óptimo en un sentido global. Carlson y Rosenblatt (1985) sugieren que la mayoría de los procedimientos del balanceo de la línea de ensamblaje se basan en un objetivo incorrecto. Los autores afirman que maximizar la ganancia (en vez de minimizar el tiempo muerto) daría una solución diferente a la mayoría de los problemas de balanceo de línea de ensamblaje, y ellos presentan varios modelos en los cuales tanto los números de estaciones como el tiempo de ciclo son variables de decisión.

Problemas para la sección 8.10

26. Considere el ejemplo de las computadoras sin marca presentado en esta sección.
    a) ¿Cuál es el tiempo de ciclo mínimo que es posible? ¿Cuál es el número mínimo de estaciones que se requerirían teóricamente para alcanzar este tiempo de ciclo?
    b) Basándose en la técnica de la jerarquización del peso posicional, ¿cuántas estaciones realmente se necesitan para el tiempo de ciclo indicado en la parte *a*?
    c) Suponga que el dueño de la compañía que vende computadoras sin marca se da cuenta de que está recibiendo órdenes de aproximadamente 100 computadoras por día. ¿Cuántas líneas de ensamblaje separadas se requieren si se supone (*i*) el mejor balance de cinco estaciones, (*ii*) el mejor balance de seis estaciones (ambos determinados en el libro), y (*iii*) el balance que usted obtuvo en el inciso *b*? Discuta los intercambios que intervienen en cada selección.
27. Una instalación de producción ensambla teléfonos baratos en una línea de producción. El ensamble requiere 15 tareas con relaciones de precedencia y tiempos de actividad como se muestran en la figura 8-16. Los tiempos de actividad aparecen junto a los números en los nodos en la red.

FIGURA 8-16
(Para el problema 27).

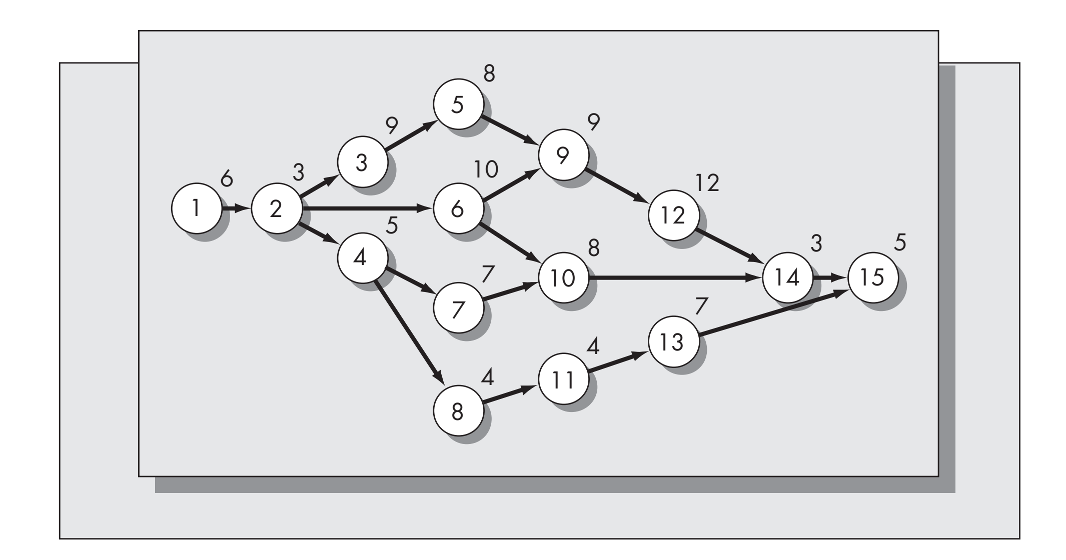
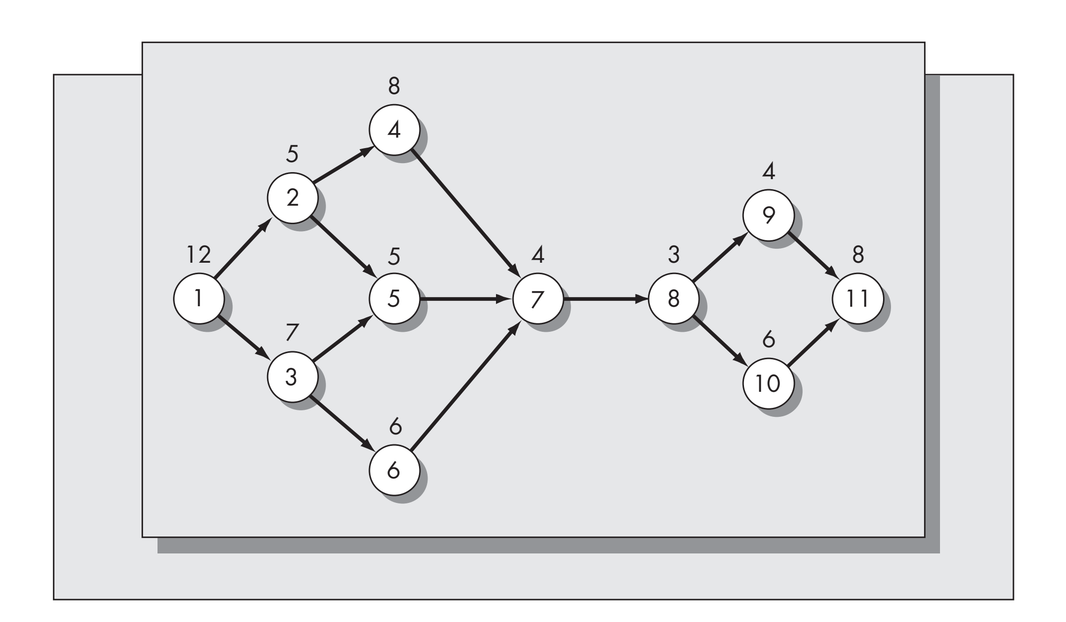


FIGURA 8-17
(Para el problema 29).



a) Determine los pesos posicionales para cada una de las actividades.

b) Para un tiempo de ciclo de 30 unidades, ¿cuál es el número mínimo de estaciones que podría alcanzarse? Encuentre el balance para $C = 30$ obtenido mediante el uso de la técnica de jerarquización del peso posicional.

c) ¿Existe una solución con el mismo número de estaciones que encontró en el inciso *b* pero con un tiempo de ciclo menor? En especial, ¿cuál resulta el tiempo de ciclo mínimo que da un balance con el mismo número de estaciones que encontró en el inciso *b*?

28. Para los datos proporcionados en el problema 27, determine por experimentación el tiempo de ciclo mínimo para un balance de tres estaciones.

29. Considere el problema del balance de la línea de ensamble representado por la red de la figura 8-17. Los tiempos de ejecución se muestran arriba de los nodos.

a) Determine un balance para $C = 15$.

b) Determine un balance para $C = 20$.

Aplicación instantánea

EL SECTOR MANUFACTURERO AHORRA CON UN SOFTWARE DE PROGRAMACIÓN

Motorola va por MOOPI

El corporativo Motorola ha optado por el MOOPI del Grupo Berclain de Canadá para programar la remanufactura de 1 000 controladores de motor cada mes en el grupo automotriz y de electrónica industrial en su planta de Seguin, Texas. La utilización de controladores usados para motor ha mejorado del 35 a más del 70% y el tiempo requerido para la programación se ha reducido de 100 a solamente ocho horas por mes como consecuencia de la implementación del software. El problema de programar la remanufactura de núcleos es muy complejo, ya que requiere cinco o seis empleados cuando se hace manualmente. De acuerdo con Eileen Svoboda, un gerente de mejora del proceso:

> La remanufactura por sí misma es un proceso único y complejo porque la materia prima —el núcleo— llega aleatoriamente con el tiempo, y el material requerido para remanufacturarla se desconoce hasta después de inspeccionar el núcleo. Ahora mismo en nuestros negocios tenemos 7 000 números de modelo de núcleo Motorola que pueden remanufacturarse como aproximadamente 800 modelos diferentes hechos sobre pedido. Para mantener todo esto en orden en el sistema manual antiguo dependíamos de senderos de papel, así como la memoria y la pericia de unas cuantas personas clave en Texas, Michigan e Illinois.

El software permite que el grupo programe un mes de producción en detalle, así como algunas funciones de planeación adicionales para un horizonte de tiempo de 26 semanas. Motorola consideró cuatro sistemas de programación diferentes. Se escogió MOOPI porque podía manejar tanto necesidades de asignación de material complejo como simulaciones fáciles de correr del tipo "que pasa si...". Aunque el aspecto de simulación jugó un papel importante en la decisión, MOOPI es básicamente un paquete de optimización y, según Berclain, la meta principal de Motorola fue optimizar el uso de los materiales.

La base de datos del sistema incorpora datos de manufactura, incluyendo la secuencia de las operaciones o ruteos, facturas de materiales e información sobre el establecimiento de secuencias. Diferencia las materias primas, el subensamblaje y los niveles del inventario de bienes terminados. Puede intercambiar datos con otros sistemas de control de inventarios, y detalla las órdenes de producción para todos los centros de trabajo.

H. P. implementa FastMAN para la programación de los materiales

Hewlett-Packard de Palo Alto, California, ha experimentado un enorme crecimiento en el negocio altamente competitivo de las PC. Se ha reportado que es la segunda compañía de computadoras personales con más rápido crecimiento en el mundo. Parte de este éxito se atribuye al manejo cuidadoso de los materiales en sus plantas de producción.

El sistema de programación de H. P. está basado en las PC y se usa conjuntamente con su sistema MRP II. Un programador post-MRP que usa H. P. es FastMAN, producido por la compañía del mismo nombre, con base en Chicago, Illinois. Actualmente, H. P. tiene ocho instalaciones de FastMAN y tres con APS, un sistema basado en estaciones de trabajo recientemente disponible con las PC. Según el doctor Lee Seaton, un veterano analista industrial con H. P.:

> El contenido de materiales de una computadora personal constituye más del 90 por ciento de su costo. Desafortunadamente, con frecuencia teníamos demasiado de una cosa y muy poco de algo más. Y sabíamos que solamente iba a empeorar, ya que los ciclos de vida del producto siguen acelerándose, recientemente descendiendo de 15 meses a menos de un año. Debimos reducir las penalizaciones de los productos y tener un corte más suave para un conjunto dado de productos. Con demasiada frecuencia lanzábamos un producto de éxito, pero no teníamos suficiente para vender o al final del ciclo sufríamos penalizaciones excesivas en los materiales.

A pesar de que el control del material es normalmente el dominio de MRP II, H. P. no estaba teniendo el nivel deseado de éxito con MRP II. El doctor Seaton afirma que el 30% de las corridas con MRP se consideraban como inútiles. Las actualizaciones de los niveles de inventario o de las facturas de materiales no se ingresaban en el sistema de manera oportuna, haciendo que los resultados fueran tardíos. Las nuevas herramientas de programación permitían a los empleados de H. P. verificar la validez de las corridas de MRP con sistemas de escritorio. También, el software suministraba la capacidad para determinar las consecuencias de las decisiones de negocios en una hoja de cálculo o en forma gráfica. Los cambios de ingeniería podían incorporarse con más facilidad en el plan de materiales, y era posible utilizar el material sobrante con mayor efectividad mediante la construcción de "conjuntos coordinados". La clave del éxito de H. P. era "llevar información a las manos de quienes tomaban las decisiones", dice Seaton.

Fuente: Estas dos aplicaciones se estudian en Parker (1995).

c) ¿Cuál resulta ser el tiempo de ciclo mínimo que puede alcanzarse empleando el número de estaciones que usted obtuvo en el inciso *a*?

d) ¿Cuál resulta ser el tiempo de ciclo mínimo que puede alcanzarse empleando el número de estaciones que usted obtuvo en el inciso *b*?

8.11 SIMULACIÓN: UNA VALIOSA HERRAMIENTA DE PROGRAMACIÓN

Frecuentemente los problemas de programación en el mundo real son demasiado complejos como para ser susceptibles al análisis matemático. La simulación basada en las computadoras es una herramienta valiosa para comparar diferentes estrategias y escenarios de programación. Algunas de las primeras aplicaciones de la investigación de operaciones basada en las computadoras incluían las simulaciones.

Una simulación es un modelo o una recreación de una situación real que permite al usuario examinar diferentes escenarios en un ambiente de laboratorio. Por ejemplo, los túneles de viento simulan condiciones de vuelo de los aeroplanos, permitiendo que un diseñador de aeroplanos pruebe diferentes configuraciones de ala sin que los aeroplanos se estrellen. Los simuladores de vuelo son valiosas herramientas en la enseñanza. Los estudiantes pueden cometer errores "fatales" en un simulador y vivir para contarlo.

Éstos son los simuladores físicos. Las aplicaciones de la planeación de la producción utilizan simuladores basados en computadoras. Una simulación mediante computadora es un programa de computadora que refleja con exactitud una situación del mundo real. Al igual que con un modelo matemático, se definen variables que representan cantidades reales y se desarrollan expresiones para describir las relaciones entre estas variables. Aunque tanto los problemas deterministas como los estocásticos son susceptibles a la simulación, se ha aplicado más este método a problemas con algún elemento de aleatoriedad. Los simuladores son especialmente valiosos para la modelación de la incertidumbre. Los científicos de la computación han invertido un gran esfuerzo en desarrollar los así llamados generadores de números aleatorios. Éstos son recursos matemáticos que crean cadenas de números que resultan ser completamente aleatorias (es decir, se extraen de una distribución uniforme). Los números aleatorios pueden transformarse en observaciones de variables aleatorias con casi cualquier distribución deseada que use relaciones de la teoría de la probabilidad. Las simulaciones que incluyen algún elemento de incertidumbre se denominan simulaciones de Monte Carlo. (Véase, por ejemplo, Fishman, 1973, para un excelente resumen de los conceptos y los métodos que se usan para construir programas de simulación y analizar la salida de estos programas.)

La simulación como herramienta de modelación está adquiriendo popularidad. Las hojas de cálculo incluyen generadores de números aleatorios que permiten construir simuladores simples. Pueden crearse modelos de hojas de cálculo más sofisticados con los productos aditivos de tercera parte (tales como @Risk de Lotus 1-2-3 y Excel). Aun cuando la mayoría de los programas de simulación pueden escribirse con facilidad directamente en un lenguaje fuente (tal como Basic o C), se dispone de muchos paquetes de simulación reservados. Ejemplos de los primeros productos son SLAM y GPSS. Desde la década de los sesenta, Pritsker Associates de West Lafayette, Indiana, ha estado comercializando simuladores basados en red diseñados específicamente para aplicaciones de la planeación de producción. Uno de los simuladores de uso general más popular disponible para la PC es SIMAN. Sin embargo, el mercado se está expandiendo rápidamente. En una encuesta de paquetes de simulación, Swain (1993) comparó 55 paquetes diferentes de software de simulación, todos diseñados para PC y estaciones de trabajo.

La mayor parte del software de programación estudiado en la sección 8.12 incorpora un componente de simulación. Abundan las aplicaciones de programación basadas en la simulación del mundo real. Por ejemplo, Hutchison y otros (1993) usaban la simulación basada en computadoras para analizar reglas de despacho para engranes en la Jeffrey Mining Equipment Division of Dresser Industries. Encontraron que la programación EDD acoplada con una planeación flexible de la capacidad condujo a una mejora potencial de cerca del 40 por ciento bajo condiciones de alta demanda. En otra aplicación, SGS Thomson aplicó la simulación para la planeación de la capacidad finita y la programación de sus áreas de la prueba y el terminado de semiconductores (Nichols, 1994). En otro estudio, se utilizó la simulación para comparar tres

políticas de programación para un sistema unitario de carga automática de almacenaje y recuperación con el objetivo de maximizar la producción total (Randhawa y Shroff, 1995). Ésta es solamente una pequeña muestra de la vasta literatura de las aplicaciones basadas en la simulación para la programación.

8.12 SOFTWARE DE PROGRAMACIÓN DE LA PRODUCCIÓN POST MRP

Aun cuando el mercado mundial de los sistemas integrados MRP es enorme (cerca de $10 billones de dólares), existe un creciente mercado para el software de programación diseñado específicamente para funcionar donde fallan los sistemas MRP. Según Parker (1995):

> Los sistemas de programación de capacidad finita aparecieron en escena porque los sistemas MRP II suponen una capacidad infinita. Los sistemas de planeación que pueden trabajar al lado de MRP II o incluso reemplazar a los módulos clave de MRP II, optimizan tanto las restricciones de materiales como de capacidad. Pero las herramientas más recientes de planeación y de programación ofrecen dos cosas adicionales. La primera es que son dinámicas, ya que su objetivo es enfrentar el cambio continuo. La segunda sería el soporte para tomar decisiones.

Parker (1994) estimó que el mercado doméstico para estas herramientas era de aproximadamente $100 millones y crecería rápidamente. Aun cuando los sistemas MRP desarrollados se basan generalmente en sistemas *mainframe* o en cliente-servidor, estos programadores post-MRP (que también se usan como sistemas individuales) se diseñan generalmente para correr en PC o estaciones de trabajo. La mayoría de los sistemas ofrecen alguna combinación de optimización para determinar horarios y simulación de Monte Carlo para permitir que el usuario explore los escenarios “qué pasa si”.

Una muestra de proveedores de estos productos incluye i2 Technologies, Dallas, Texas; The Berclain Group, Quebec City; Red Pepper Software, San Mateo, California; Bridgeware, Hayward, California; FastMAN, Chicago, Illinois; Orissa International, Torrance, California; y Enterprise Planning Systems, Wheaton, Illinois. Varios de los suministradores del software MRP estudiado en el capítulo 7 también están incorporando algunas de las características de estos programadores, incluyendo apoyo para simulación y decisión, directamente en los sistemas MRP. La serie Preactor de software de programación de piso de Systems Modeling Corporation (Vasilash, 1995) es nueva en el mercado.

Una diferencia importante entre los planificadores y los programadores es la longitud del horizonte de planeación. De acuerdo con Sanjeev Gupta, presidente de Orissa International, “la planeación se hace para un periodo de seis a nueve meses, mientras que la programación de capacidad finita examina un periodo de dos a tres meses, aunque depende de cual sea el tiempo de ciclo del producto. El despacho, que está sujeto a la información en tiempo real, se hace en un periodo de dos a tres días”.

Aunque el valor de la tecnología de productos derivados de otras investigaciones proveniente del gasto militar está siendo objeto de un profundo escrutinio, la tecnología de los programadores de capacidad finita puede beneficiarse con el trabajo hecho en la NASA para los sistemas espaciales de programación. El Advanced Scheduling Environment, que actualmente comercializa Avyx, Inc., de Englewood, Colorado, posibilita describir genéricamente un amplio rango de aplicaciones de programación. El sistema está constituido por un conjunto de estructuras de datos que alojan una descripción muy general de los recursos, las actividades y las restricciones de una fábrica. El software usa estructuras de datos tipo árbol para resolver problemas de programación, en lugar de las bases de datos relacionales mas tradicionales que caracterizan a la gran mayoría de los sistemas de programación y MRP.

8.13 NOTAS HISTÓRICAS

El interés en el estudio de los efectos de diferentes estrategias de secuenciación en el taller de trabajo es relativamente reciente. Una de las monografías pioneras en la literatura abierta que consideró los aspectos de la secuenciación es la de Salveson (1952). El primer trabajo publicado

importante que presentó resultados analíticos de las estrategias óptimas de secuenciación es el de Johnson (1954). El artículo de revisión de Bellman (1956) discute la programación de secuencias (entre otros temas) y presenta una prueba interesante de la optimalidad del algoritmo de Johnson basada en argumentos de programación dinámica. Bellman parecería ser el primero en reconocer la optimalidad de un programa de permutación para programar n trabajos en tres máquinas.

El trabajo de Johnson despertó un interés considerable en la programación. La excelente monografía de Conway, Maxwell y Miller (1967), por ejemplo, lista más de 200 publicaciones hasta 1967. Es probable que desde esa época se hayan publicado mucho más de 200 artículos. Gran parte de la investigación reciente en la programación de secuencias se ha centrado en la programación estocástica. Weiss (1982) suministra una interesante sinopsis de su trabajo en esta área. Es posible que Bruno y Downey (1977) y Pinedo y Weiss (1979) descubrieran casi al mismo tiempo que los trabajos deben programarse en orden de tiempo decreciente de procesamiento esperado si la programación se hace en dos procesadores paralelos y el objetivo es minimizar la terminación esperada. El modelo elegante de prueba que presentamos es de Pinedo y Weiss (1979).

Gran parte del trabajo sobre programación dinámica bajo la incertidumbre aparece en la monografía de Conway, Maxwell y Miller (1967). La regla $c\mu$, evidentemente descubierta por Derman y otros (1978) en un contexto completamente diferente, condujo a varias ampliaciones. Véanse, por ejemplo, Righter y Shanthikumar (1989).

8.14 Resumen

En muchos contextos de las operaciones de administración aparecen problemas de programación. Una de las áreas exploradas en este capítulo fueron las reglas óptimas de secuenciación para programar trabajos en máquinas, un problema que surge en el *control de piso de almacén*. También consideramos los problemas de programación dinámica, que son más frecuentes en los tipos de problemas de programación que se encuentran en la administración de los sistemas de servicios.

Gran parte del énfasis del capítulo estuvo en la determinar *reglas de secuenciación* eficientes. La forma de la regla de secuenciación óptima obedece a varios factores, incluyendo el patrón de llegada de los trabajos, la configuración del taller de tareas, las restricciones y los objetivos de la optimización. Los objetivos comunes en la administración del taller de tareas incluyen el cumplimiento de las fechas de vencimiento, la minimización del trabajo en proceso, la minimización del tiempo de flujo promedio y la minimización del tiempo muerto de máquina.

Estudiamos cuatro reglas de secuenciación: FCFS (primeras llegadas, primeras salidas), SPT (tiempo de procesamiento más corto primero), EDD (primera fecha de entrega) y CR (razón crítica). Para trabajos de secuenciación en una sola máquina, demostramos que SPT optimizaba varios objetivos, incluyendo el tiempo de flujo medio. Sin embargo, la secuenciación SPT podría ser problemática. Los trabajos largos serían empujados constantemente hasta el final de la cola de trabajos y tal vez nunca llegarían a ser procesados. Por esa razón, la programación SPT pura se usa rara vez en la práctica. La *programación basada en la razón crítica* intenta balancear la importancia puesta en el tiempo de procesamiento y en el tiempo restante hasta la fecha de vencimiento. Sin embargo, hay poca evidencia que sugiera que la programación de la razón crítica se desempeñe bien en relación con los criterios comunes de optimización tales como el tiempo de flujo medio. Según podría esperarse, la programación de la *primera fecha de entrega* es la que tiene el mejor desempeño cuando el objetivo es minimizar el retardo máximo.

Consideramos varios algoritmos para la secuenciación de una sola máquina, incluyendo el algoritmo de Moore (1968) para minimizar el número de trabajos retardados y el algoritmo de Lawler (1973) para minimizar cualquier función no decreciente del flujo sujeto a las restricciones de precedencia. Un resumen excelente de los algoritmos para la programación de secuencias puede encontrarse en French (1982). En este capítulo también se presentaron varias técnicas para la programación *de máquinas múltiples*. Se estudia el trabajo clásico de Johnson (1954) sobre la programación de *n trabajos a través de dos máquinas*, así como el método gráfico de Akers (1956) para *dos trabajos a través de m máquinas*.

Los problemas de programación *estocástica* se trataron en dos contextos: el estático y el dinámico. Los problemas estáticos son la contraparte estocástica de los problemas tratados anteriormente en el capítulo, excepto que los tiempos de terminación de los trabajos son variables aleatorias. La mayoría de los resultados simples para los problemas estáticos de programación estocástica requieren partir de la hipótesis de que los tiempos de los trabajos tienen una distribución exponencial. Debido a la propiedad sin memoria de la distribución exponencial, probablemente esta hipótesis no se satisface en la mayoría de los talleres de trabajo de manufactura. El caso dinámico se presenta cuando los trabajos llegan aleatoriamente con el tiempo. La *teoría de colas*, estudiada con detalle en el suplemento 2, es el medio para analizar estos problemas.

El *problema de balanceo de la línea de ensamble* es aquél en el cual debe ejecutarse un grupo de tareas para cada artículo. Además, las tareas deben ser realizadas en una secuencia específica. El reto consiste en asignar las tareas a las estaciones de trabajo en una línea de ensamblaje. La calidad del balance se mide por el tiempo muerto que resulta en cada estación. Determinar la mejor combinación de tiempos de ciclo (el lapso asignado a cada estación) y el número de estaciones es un problema analítico muy difícil. Estudiamos un método de solución heurística muy simple de Helgeson y Birnie (1961), conocido como la técnica de jerarquización de los pesos posicionales, que suministra rápidamente un balance eficiente.

Aun cuando la mayor parte de este capítulo trata sobre métodos analíticos para el desarrollo de horarios, es importante tener en mente que muchos problemas reales de programación resultan demasiado complejos como para modelarse matemáticamente. En estos casos, la *simulación* puede ser una herramienta muy valiosa. Una simulación es un modelo de un proceso real expresado en un programa de computadora. Al correr simulaciones sobre diferentes escenarios, las consecuencias de las estrategias alternativas pueden evaluarse fácilmente. Las simulaciones son efectivas especialmente cuando existen aleatoriedad o variaciones significativas.

El *software de programación* para el taller es un negocio floreciente. Los programas diseñados para correr en las PC y en estaciones de trabajo están disponibles con diferentes vendedores. La mayor parte de estos programas suministran interfaces convenientes para los sistemas existentes MRP. Aunque el mercado para los programadores post MRP es mucho más pequeño que el de los sistemas desarrollados integrados MRP II, está creciendo y probablemente sobrepasa los $100 millones solamente en los Estados Unidos. Estos programas se diseñan para correr conjuntamente con un sistema MRP, y generalmente incorporan alguna mezcla de optimización y simulación. Muchas compañías han tenido éxito al implementar estos programas en el piso de fábrica.

Existen varios libros excelentes en torno a la programación. Para lecturas adicionales sobre esta área, sugeriríamos los textos de Baker (1974); French (1982); y Conway, Maxwell y Miller (1967). Pinedo (1995) suministra la cobertura más actualizada y detallada.

Problemas adicionales sobre pronóstico

30. El taller de hojalatería de Mike tiene cinco automóviles en espera de ser reparados. El taller es bastante pequeño, de modo que sólo se puede reparar un automóvil a la vez. El número de días que se requieren para reparar cada automóvil y la fecha prometida para cada uno se dan en la siguiente tabla.

Automóviles	Tiempo de reparación (días)	Fecha prometida
1	3	5
2	2	6
3	1	9
4	4	11
5	5	8

Mike se comprometió a dar un automóvil en renta a cada cliente cuyo vehículo no se repare a tiempo. Compare el desempeño de las cuatro reglas de secuenciación FCFS, SPT, EDD y CR relacionándolas con la minimización del retardo promedio.

31. Para cada uno de los problemas listados, indique precisamente qué o quién correspondería a los trabajos y quién o qué correspondería a las máquinas. En cada caso discuta cuáles objetivos pueden ser adecuados, y las prioridades especiales que puedan existir.

a) Ingresar a los pacientes en la sala de emergencias de un hospital.

b) Desembarcar la carga de los barcos en el puerto.

c) Servir a los usuarios en un sistema de computadora de tiempo compartido.

d) Transferir las llamadas telefónicas de larga distancia de una ciudad a otra.

32. Seis pacientes están esperando en una sala de emergencia de un hospital para recibir cuidados por padecimientos varios, que varían desde un esguince de tobillo hasta una herida por arma de fuego. Los pacientes están numerados en la secuencia de llegada y se les da un peso por prioridad basado en la gravedad de su dolencia. Sólo hay un doctor disponible.

Paciente	1	2	3	4	5	6
Tiempo requerido	20 min	2 h	30 min	10 min	40 min	1 h
Prioridad	1	10	3	5	2	2

a) Suponga que los pacientes se programan sobre la base de FCFS. Calcule el tiempo de flujo medio y el tiempo de flujo medio ponderado, donde el peso es el número de prioridad asociado a cada paciente.

b) Realice los cálculos del inciso *a* para la secuencia de SPT.

c) Determine una secuencia diferente a las de los incisos *a* y *b* para obtener un valor menor del tiempo de flujo medio ponderado.

33. Considere el taller de hojalatería de Mike mencionado en el problema 30. ¿Qué secuencia de los trabajos minimiza:

a) el tiempo de flujo medio?

b) el retardo máximo?

c) el número de trabajos retardados?

34. Considere la situación de la sala de emergencias mencionada en el problema 32. Determine la secuencia de tratamiento de los pacientes con objeto de minimizar el valor *ponderado* del tiempo de flujo medio.

35. La compañía de helados de Barbara y Jenny produce cuatro diferentes sabores de helado: vainilla, chocolate, fresa y dulce de cacahuate. Cada lote de helado se produce en la misma cuba grande, que debe limpiarse antes de cambiar los sabores. Se requiere un día para el proceso de limpieza.

En el momento presente se tienen las siguientes órdenes pendientes de helado:

Sabor	**Tamaño de la orden (galones)**	**Fecha de entrega**
Vanilla	385	3
Chocolate	440	8
Fresa	200	6
Dulce de cacahuate	180	12

Toma un día producir el helado y puede producirse un máximo de 100 galones a la vez. Se requiere limpieza solamente cuando se cambian sabores. Debe terminarse la producción de un sabor antes de iniciar la de otro. La limpieza siempre se inicia al comenzar el día.

Si se trata a cada sabor de helado como un trabajo distinto, encuentre lo siguiente:

a) La secuencia de producción de los sabores de helado con objeto de minimizar el tiempo de flujo medio para todos los sabores.

b) La secuencia óptima de producción de los sabores con objeto de minimizar el número de sabores retrasados.

36. Considere a la compañía de helados de Barbara y Jenny, mencionada en el problema 35. Suponga que si se producen vainilla o fresa después de chocolate o dulce de cacahuate, se requiere un día extra de limpieza. Por esa razón deciden que los helados de vainilla y de fresa van a producirse antes que los de chocolate y de dulce de cacahuate.

a) Encuentre la secuencia óptima de los sabores para minimizar el retraso máximo usando el algoritmo de Lawler.

b) Enumere todas las secuencias factibles y determine la secuencia que minimice el retraso máximo al evaluar y comparar el valor de la función objetivo para cada caso.

37. Irving Bonner, un consultor independiente de programación de computadoras, ha obtenido un contrato para terminar ocho trabajos de programación de computadoras. Algunos trabajos deben terminarse con cierta secuencia porque incluyen módulos de programa que van a enlazarse.

Trabajo	Tiempo requerido (días)	Fecha de entrega
1	4	Junio 8
2	10	Junio 15
3	2	Junio 10
4	1	Junio 12
5	8	Julio 1
6	3	Julio 6
7	2	Junio 25
8	6	Junio 29

Restricciones de precedencia:

$$1 \rightarrow 2 \rightarrow 5 \rightarrow 6.$$
$$4 \rightarrow 7 \rightarrow 8.$$

Suponga que la fecha actual es lunes 1 de junio y que Bonner no trabaja los fines de semana. Usando el algoritmo de Lawler, encuentre la secuencia de ejecución de los trabajos para minimizar el retraso máximo sujeto a las restricciones de precedencia.

38. William Beebe es dueño de una pequeña zapatería. Recibe 10 pares de zapatos que requieren suelas nuevas y boleo. Cuenta con una máquina que puede poner las suelas a un par de zapatos a la vez, y el tiempo requerido para la operación varía según el tipo y la condición del zapato, y el tipo de suela que se use. Los zapatos se bolean en una máquina dedicada también a este objetivo, ya que el boleo siempre se hace después de poner las suelas.

Generalmente su ayudante bolea mientras el señor Beebe pone las suelas. Los tiempos que tardan colocar suelas y boleo son (en minutos)

Zapatos	Tiempo de colocación de suelas	Tiempo de boleo
1	14	3
2	28	1
3	12	2
4	6	5
5	10	10
6	14	6
7	4	12
8	25	8
9	15	5
10	10	5

¿En qué orden deben repararse los zapatos con objeto de minimizar la terminación total de estos 10 trabajos?

39. Una fábrica de artículos de cuero tiene dos prensas para troquelar orificios en los diferentes artículos producidos antes de la costura. Suponga que deben procesarse 12 trabajos diferentes en una o en otra prensa troqueladora (es decir, procesamiento paralelo). Los tiempos de procesamiento (en minutos) para estos 12 trabajos se dan en la siguiente tabla.

Trabajo	1	2	3	4	5	6	7	8	9	10	11	12
Tiempo	26	12	8	42	35	30	29	21	25	15	4	75

Suponga que ambas prensas están inicialmente inactivas. Compare el desempeño de las reglas de SPT y de LPT (el tiempo de procesamiento más largo) para este ejemplo. (En el procesamiento paralelo el siguiente trabajo simplemente se programan para la otra máquina disponible.)

40. Una contadora independiente planea preparar declaraciones de impuestos para seis de sus clientes. Antes de la preparación real de cada declaración, su secretaria revisa el archivo del cliente para asegurarse de que toda la documentación necesaria esté ahí y obtiene todos los formatos necesarios al respecto. Basándose en la experiencia pasada con los clientes, su secretaria estima que se requieren los siguientes tiempos (en horas) para preparar la declaración y para que la contadora termine el papeleo necesario antes de entregarla:

Cliente	Tiempo de secretaria	Tiempo de contadora
1	1.2	2.5
2	1.6	4.5
3	2.0	2.0
4	1.5	6.0
5	3.1	5.0
6	0.5	1.5

¿En qué orden debe terminarse el trabajo con miras a reducir el tiempo total requerido para los seis clientes?

41. Cinco sastres de Hong Kong: Simon, Pat, Choon, Paul y Wu, deben terminar las modificaciones de un traje para el duque y un vestido para la duquesa tan rápidamente como sea posible. Para el vestido, Choon debe pasar primero 45 minutos cortando la tela, luego Pat va a ocupar 75 minutos cosiendo el corpiño, Simon va a necesitar 30 minutos basteando las

mangas, Paul 2 horas sujetando con cordones la bastilla y finalmente Wu necesitará 80 minutos para terminar los detalles. Por lo que toca al traje, Pat comienza acortando las mangas, lo que le lleva 100 minutos. Entonces sigue Paul, quien cose el forro en 1.75 horas; a continuación Wu ocupa 90 minutos pegando los botones y angostando las solapas; y, finalmente, Choon plancha y limpia el traje en 30 minutos.

Determine con precisión cuándo debe ejecutar cada sastre la tarea que le corresponde con objeto de minimizar el tiempo total requerido para terminar el vestido y el traje. Suponga que comienzan a trabajar a las 9 a. m. y no paran. Dibuje el diagrama de Gantt que indique su solución.

42. Ensamblar un radio reloj transistorizado requiere un total de 11 tareas. Los tiempos de las tareas y las relaciones de los predecesores se dan en la siguiente tabla.

Tarea	**Tiempo (segundos)**	**Predecesores inmediatos**
1	4	
2	38	
3	45	
4	12	1, 2
5	10	2
6	8	4
7	12	5
8	10	6
9	2	7
10	10	8, 9
11	34	3, 10

a) Desarrolle una red para esta operación de ensamblado.

b) ¿Cuál es el tiempo de ciclo mínimo que podría considerarse para esta operación? ¿Cuál es el número mínimo de estaciones que podrían usarse con este tiempo de ciclo?

c) Empleando la técnica de jerarquización del peso posicional, determine el balance que resulta al usar un tiempo de ciclo de 45 segundos.

d) Determine por experimentación el tiempo de ciclo mínimo que conduce a un balance de cuatro estaciones.

e) ¿Cuál es la tasa diaria de producción para este producto si la compañía adopta el balance que usted determinó en el inciso *c*? (Suponga un día de seis horas para sus cálculos.) ¿Qué tendría que hacerse si la compañía quisiera una tasa de producción más alta?

43. Suponga que en el problema 42 surgen restricciones adicionales del hecho de que ciertas tareas no pueden ejecutarse en la misma estación. En especial piense que las tareas se dividirían en zonas de la siguiente manera:

Zona 1	Tareas 2, 3, 1, 4, 6
Zona 2	Tareas 5, 8, 7, 9
Zona 3	Tareas 10, 11

Suponiendo que sólo pueden ejecutarse tareas en la misma categoría de zona para la misma estación, determine el balance de línea resultante para el problema 42 basado en un tiempo de ciclo de 45 segundos.

44. La Southeastern Sports Company produce palos de golf en una línea de ensamble en su planta de Marietta, Georgia. El ensamblaje final de las maderas requiere las ocho operaciones dadas en la siguiente tabla.

Tarea	Tiempo requerido (min.)	Predecesores inmediatos
1. Pulir el vástago	12	
2. Esmerilar el extremo del vástago	14	
3. Pulir el extremo del palo de golf	6	
4. Grabar el número	4	3
5. Conectar la madera al vástago	6	1, 2, 4
6. Colocar y asegurar el broche de conexión	3	5
7. Poner pegamento en el otro extremo del vástago	3	1
8. Colocar el mango y balancear	12	6, 7

a) Dibuje una red para representar la operación de ensamblado.

b) ¿Cuál es el tiempo de ciclo mínimo que puede considerarse? Determine el balance que resulta aplicando la técnica de jerarquización de peso posicional para este tiempo de ciclo.

c) Por experimentación, determine el tiempo de ciclo mínimo que puede alcanzarse con un balance de tres estaciones.

45. Desarrolle una plantilla que calcule varias medidas de desempeño para una secuenciación de trabajos de primeras llegadas, primeras salidas. Permita hasta 20 trabajos de modo que la columna 1 contenga los números 1, 2, . . . , 20. La columna 2 debe listar los tiempos de procesamiento que el usuario va a ingresar y la columna 3 las fechas de vencimiento que también va a ingresar el usuario. La columna 4 registra el retardo y, la columna 5, el tiempo de flujo. Desarrolle la lógica para calcular el tiempo de flujo medio, el retardo promedio y el número de trabajos retardados. (Cuando se calcule el promedio de una columna, asegúrese de que su hoja de cálculo no trate a los espacios vacíos como ceros.)

a) Use su plantilla para encontrar el tiempo de flujo medio, el retardo promedio y el número de trabajos retardados para el problema 7 (sección 8.6), suponiendo una secuencia FCFS.

b) Encuentre el tiempo de flujo medio, el retardo promedio, y el número de trabajos retardados para el problema 8, suponiendo una secuencia FCFS.

c) Encuentre el tiempo de flujo medio, el retardo promedio y el número de trabajos retardados suponiendo una secuencia FCFS para el siguiente problema de 20 trabajos:

Tarea	Tiempo de procesamiento	Fecha de entrega	Tarea	Tiempo de procesamiento	Fecha de entrega
1	10	34	11	17	140
2	24	38	12	8	120
3	16	60	13	23	110
4	8	52	14	25	160
5	14	25	15	40	180
6	19	95	16	19	140
7	26	92	17	6	130
8	24	61	18	23	190
9	4	42	19	25	220
10	12	170	20	14	110

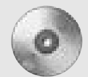


46. *a*) Resuelva el problema 45*a*) suponiendo una secuencia SPT. Puede emplear la hoja de cálculo desarrollada en el problema 45 y simplemente ordenar los datos en las primeras

tres columnas, usando la columna 2 (el tiempo de procesamiento) como una clave de clasificación.

b) Resuelva el problema 45*b*) suponiendo una secuencia SPT.

c) Resuelva el problema 45*c*) suponiendo una secuencia SPT.



47. *a*) Resuelva el problema 45*a*) suponiendo una secuencia EDD. En este caso se hace la clasificación en la columna de la fecha de entrega.

b) Resuelva el problema 45*b*) suponiendo una secuencia EDD.

c) Resuelva el problema 45*c*) suponiendo una secuencia EDD.

Bibliografía

Akers, S. B. "A Graphical Approach to Production Scheduling Problems". *Operations Research* 4 (1956), pp. 244-245.

Arcus, A. L. "COMSOAL: A Computer Method for Sequencing Operations for Assembly Lines". *International Journal of Production Research* 4, 1966, pp. 259-277.

Baker, K. R. *Introduction to Sequencing and Scheduling.* Nueva York: John Wiley & Sons, 1974.

Banerjee, B. P. "Single Facility Sequencing with Random Execution Times". *Operations Research* 13, 1965, pp. 358-364.

Bellman, R. E. "Mathematical Aspects of Scheduling Theory". *SIAM Journal of Applied Mathematics* 4, 1956, pp. 168-205.

Bruno, J. y P. Downey. "Sequencing Tasks with Exponential Service Times on Two Machines". Technical Report, Department of Electrical Engineering and Computer Science, University of California Santa Bárbara, 1977.

Carlson, R. y M. Rosenblatt. "Designing a Production Line to Maximize Profit". *IIE Transactions* 17, 1985, pp. 117-122.

Conway, R. W., W. L. Maxwell y L. W. Miller. *Theory of Scheduling.* Reading, MA: Addison Wesley, 1967.

Derman, C., G. L. Lieberman y S. M. Ross. "A Renewal Decision Problem". *Management Science* 24, 1978, pp. 554-561.

Eilon, S., C. D. T. Watson-Gandy y N. Christofides. *Distribution Management: Mathematical Modeling and Practical Analysis.* Londres: Griffin, 1971.

Fisher, M., A. J. Greenfield, R. Jaikumar y J. T. Uster III. "A Computerized Vehicle Routing Application". *Interfaces* 12, 1982, pp. 42-52.

Fishman, G. S. *Concepts and Methods in Discrete Event Digital Simulation.* Nueva York: John Wiley & sons, 1973.

French, S. *Sequencing and Scheduling: An Introduction to the Mathematics of the Job Shop.* Chichester, Inglaterra: Ellis Horwood Limited, 1982.

Graves, S. C. "A Review of Production Scheduling". *Operations Research* 29, 1981, pp. 646-675.

Held, M., R. M. Karp y R. Shareshian. "Assembly Line Balancing—Dynamic Programming with Precedence Constraints". *Operations Research* 11, 1963, pp. 442-459.

Helgeson, W. P. y D. P. Birnie. "Assembly Line Balancing Using the Ranked Positional Weight Technique". *Journal of Industrial Engineering* 12, 1961, pp. 394-398.

Hillier, F. S. y R. W. Boling. "On the Optimal Allocation of Work in Symmetrically Unbalanced Production Line Systems with Variable Operation Times". *Management Science* 25, 1986, pp. 721-728.

Hutchison, J., G. Leong y P. T. Ward. "Improving Delivery Performance in Gear Manufacturing at Jeffrey Division of Dresser Industries". *Interfaces* 23, núm. 2, marzo-abril 1993, pp. 69-83.

Johnson, S. M. "Optimal Two and Three Stage Production Schedules with Setup Times Included". *Naval Research Logistics Quarterly* 1, 1954, pp. 61-68.

Kilbridge, M. D. y L. Wester. "A Heuristic Method of Line Balancing". *Journal of Industrial Engineering* 12, 1961, pp. 292-298.

Lawler, E. L. "Optimal Sequencing of a Single Machine Subject to Precedence Constraints". *Management Science* 19, 1973, pp. 544-546.

Melnyk, S. A., P. L. Carter, D. M. Dilts y D. M. Lyth. *Shop Floor Control.* Homewood IL: Dow Jones-Irwin, 1985.

Moore, J. M. "An *n*-job, One Machine Sequencing Algorithm for Minimizing the Number of Late Jobs". *Management Science* 15, 1968, pp. 102-109.

Nichols, J. C. "Planning for Real World Production". *Production* 106, núm. 8, agosto 1994, pp. 18-20.

Parker, K. "What New Tools Will Best Tame Time". *Manufacturing Systems* 12, núm. 1, enero 1994, pp. 16-22.

Parker, K. "Dynamism and Decision Support". *Manufacturing Systems* 13, núm. 4, abril 1995, pp. 12-24.

Pinedo, M. "Stochastic Scheduling with Release Dates and Due Dates". *Operations Research* 31, 1983, pp. 554-572.

Pinedo, M. *Scheduling, Theory, Algorithms and Systems.* Englewood Cliffs. NJ: Prentice Hall, 1995.

Pinedo, M. y G. Weiss. "Scheduling Stochastic Tasks on Two Parallel Processors". *Naval Research Logistics Quarterly* 26, 1979, pp. 527-536.

Randhawa, S. U. y R. Shroff. "Simulation-Based Design Evaluation of Unit Load Automated Storage/Retrieval Systems". *Computers and Industrial Engineering* 28, núm. 1, enero 1995, pp. 71-79.

Righter, R. "Job Scheduling to Minimize Weighted Flowtime on Uniform Processors". *Systems and Control Letters* 10, 1988, pp. 211-216.

Righter, R. y J. G. Shanthikumar. "Scheduling Multiclass Single-Server Queuing Systems to Stochastically Maximize the Number of Successful Departures". *Probability in the Engineering and Informational Sciences* 3, 1989, pp. 323-333.

Rothkopf, M. S. "Scheduling with Random Service Times". *Management Science* 12, 1966, pp. 707-713.

Salveson, M. E. "Production Planning and Scheduling". *Econometrica* 20, 1952, pp. 554-590.

Schrage, L. "Formulation and Structure of More Complex/Realistic Routing and Scheduling Problems". *Networks* 11, 1981, pp. 229-232.

Swain, J. J. "Flexible Tools for Modeling". *OR/MS Today* 20, núm. 6, 1993, pp. 62-78.

Thangavelu, S. R. y C. M. Shetty. "Assembly Line Balancing by Zero One Integer Programming". *AIIE Transactions* 3, 1971, pp. 61-68.

Vasilash, G. "Scheduling for the Shop Floor". *Production* 107, núm. 6, junio 1995, pp. 46-47.

Vollman, T. E., W. L. Berry y D. C. Whybark. *Manufacturing Planning and Control Systems*, 3a. ed. Homewood, IL: Richard D. Irwin, 1992.

Weiss, G. "Multiserver Stochastic Scheduling". *Deterministic and Stochastic Scheduling.* Memorias de NATO Advanced Study and Research Institute on Theoretical Approaches to Scheduling Problems. Norwell, MA: D. Reidel, 1982.

Suplemento **Dos**

Teoría de colas

S2.1 INTRODUCCIÓN

La teoría de colas es el estudio de los procesos de línea de espera. Una *cola* (*queue*) es la versión británica de la línea de espera norteamericana (*waiting line*). Virtualmente todos los resultados de la teoría de colas suponen que tanto los procesos de llegada como de servicio son aleatorios. Es la interacción entre estos dos procesos lo que hace a las colas un área interesante y desafiante. En la figura S2-1 mostramos un sistema típico de colas. Los clientes llegan a una o más instalaciones de servicio. Si otras personas ya están esperando, según el tipo de servicio, los clientes que acaban de llegar esperarían su turno para ser atendidos por el siguiente servidor disponible y saldrían del sistema cuando terminen sus operaciones.

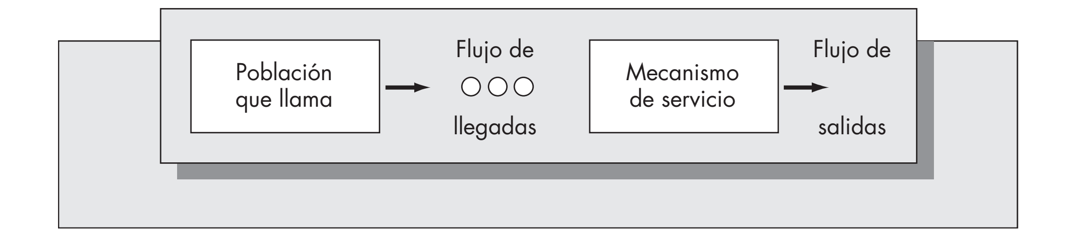
Los problemas de colas son comunes en la administración de las operaciones. En el contexto de las manufacturas, el ambiente de taller puede considerarse una compleja red interrelacionada de colas. A medida que los trabajos terminan de realizarse en un centro, éstos hacen cola para ser procesados en el centro siguiente. Si la programación en un ambiente de taller se asume como un problema dinámico en lugar de uno estático, deben hallarse las tasas de flujo y otras medidas usando los resultados de la teoría de colas.

Los problemas de colas se presentan con mucha frecuencia en los sistemas de servicios. Cada uno de nosotros experimenta colas en los sistemas de servicios todos los días. De hecho, según Michael Fortino, de Priority Management en Pittsburgh, la mayoría de las personas pasan aproximadamente cinco años de su vida haciendo colas y seis meses esperando en los semáforos (*Chicago Tribune*, 21 de junio, 1988).

Hacemos cola para esperar nuestro turno en bancos, supermercados, estéticas, casetas de pago en autopistas, y restaurantes. La programación de los aterrizajes de los aeroplanos en una pista específica es un problema de colas: piense que los aeroplanos en el aire son clientes que esperan ser atendidos, y la pista es el servidor. Los sistemas de tráfico telefónico son ejemplos de redes complejas de colas. Las llamadas telefónicas se enrutan a través de los sistemas de distribución, donde hacen cola hasta que son desviados ya sea a la siguiente estación de distribución o bien a su destino final. En efecto, fue A. K. Erlang, un ingeniero de telefonía danés, el responsable de muchos de los primeros desarrollos teóricos en el área de las colas.

Las medidas importantes del desempeño para cierta clase de colas disponen de fórmulas relativamente simples. Para esta clase, tanto los patrones de servicio como de llegada son completamente aleatorios (el significado de “completamente aleatorio” en este contexto se estudiará en la sección S2.5). Sin embargo, si el sistema presenta un comportamiento que no sea completamente aleatorio, se dificulta encontrar resultados útiles. Por esa razón, la simulación de Monte Carlo es una herramienta común para analizar problemas complejos de colas.

FIGURA S2-1
Sistema típico de colas para un solo servidor.



S2.2 ASPECTOS ESTRUCTURALES DE MODELOS DE COLAS

1. *Proceso de llegada*. Describe las llegadas de los clientes al sistema y se caracteriza por la distribución de los tiempos entre llegadas. El caso más simple es cuando se trata de una sola llegada a la vez, de manera completamente aleatoria. Esto da lugar a una distribución exponencial para los tiempos entre llegadas. Otras posibilidades incluyen las distribuciones generales entre llegadas o más de una llegada a la vez (lo que se conoce en las colas como llegadas por lote).
2. *Proceso del servicio*. Se distingue por la distribución del tiempo requerido para servir a un cliente. Nuevamente, el caso más fácil de analizar se presenta cuando la distribución de los tiempos de servicio es exponencial. Otras distribuciones dan lugar a modelos de colas más complejos.
3. *Disciplina de servicio*. Ésta es la regla mediante la cual se sirve a los clientes en la cola. La mayoría de los problemas de colas que se presentan en los sistemas de servicio son del tipo primeras llegadas, primeras salidas (FCFS). Generalmente consideramos a esta regla como "justa". Sin embargo, también son comunes otras disciplinas de servicio. Cuando compramos leche, podemos verificar las fechas en los envases de cartón y elegir la que tenga la fecha de caducidad más lejana. Si consideramos que la leche es la cola, esto significa que la disciplina de servicio es del tipo último en llegar, primero en salir (LCFS). Las salas de emergencia de los hospitales darán prioridad a los pacientes cuya vida está en peligro —tales como los traumatismos ocasionados en un accidente automovilístico— por sobre aquéllos con problemas menos graves. Esto se denominaría una disciplina de servicio por prioridad.
4. *Capacidad de la cola*. En algunos casos, el tamaño de la cola puede estar limitado. Por ejemplo, los restaurantes y las salas de cine pueden alojar solamente a un número determinado de clientes. Desde un punto de vista matemático, la hipótesis más simple es que el tamaño de la cola sea ilimitado. Aun si la capacidad es finita, es razonable ignorar dicha restricción si es poco probable que la cola se sature.
5. *Número de servidores*. Las colas pueden ser atendidas por un solo servidor o por servidores múltiples. Este último caso es común en los bancos: los clientes forman una sola cola y son atendidos por el siguiente servidor disponible. En contraste, el área de despacho de un supermercado típico *no* es una cola con servidores múltiples. Como un comprador debe formarse en una sola cola específica, éste es un sistema paralelo de colas con un solo servidor (posiblemente dependiente). Otro ejemplo de una cola con servidores múltiples sería el área de aterrizaje del aeropuerto; los aeroplanos pueden despegar o aterrizar en una de varias pistas.
6. *Estructura de la red*. Se forma una red de colas cuando la salida de una da pie a la entrada de otra cola. La mayoría de los procesos de manufactura generalmente es un tipo de red de colas. Los sistemas de carreteras y los sistemas de distribución telefónica son otros ejemplos. Estas estructuras con frecuencia son demasiado complejas como para analizarlas matemáticamente.

S2.3 NOTACIÓN

Resumen

$N(t)$ = Número de clientes en el sistema en el instante t.

$P_n(t) = P\{N(t) = n \mid N(0) = 0\}$.

P_n = Probabilidad de estado estable de n clientes en el sistema.

λ_n = Tasa de llegada cuando hay n clientes en el sistema.

μ_n = Tasa de servicio cuando hay n clientes en el sistema.

c = número de servidores.

ρ = tasa de utilización.

L = Número esperado de clientes en el sistema en estado estable.

L_q = Número esperado de clientes en la cola en estado estable.

W = Tiempo esperado de un cliente en estado estable.

W_q = Tiempo esperado en la cola en estado estable.

K = Número máximo en el sistema para un sistema con capacidad finita.

Discusión

Para los casos que consideramos, $N(t)$ es el estado del sistema para el instante t. Esto significa que, dado el conocimiento del valor presente de $N(t)$, la posibilidad de valores futuros de $N(t)$ es independiente de los valores pasados (se le denomina propiedad markoviana). Para los sistemas que vamos a examinar, la definición de $N(t)$ como el número total de clientes en el sistema para el instante t va a garantizar que $N(t)$ tiene la propiedad markoviana y clasifica como la definición del estado del sistema. El número de clientes en el sistema incluye tanto a los que esperan por el servicio (es decir, en la cola) como a los que ya se está atendiendo. Nosotros permitimos que n sea un valor fijo específico de $N(t)$.

$P_n(t)$ es la probabilidad transitoria de que haya n clientes en el sistema para el instante t, pues éste está vacío para el instante 0. Para nuestros propósitos, sólo nos interesa el estado estable o la probabilidad estacionaria de n clientes en el sistema, P_n, dada por

$$P_n = \lim_{t \to \infty} P_n(t).$$

Sea c el número de servidores. El caso $c > 1$ corresponde a los servidores múltiples en paralelo. La mayoría de nuestras aplicaciones serán para el caso $c = 1$.

La tasa de llegada, λ_n, es la tasa esperada o promedio de clientes que llegan al sistema y se expresa como el número de llegadas por unidad de tiempo. Observe que permitir que la tasa de llegada dependa de n implica que podemos considerar casos en los cuales los clientes podrían escoger no formarse en la cola basándose en su tamaño. Generalmente, suponemos $\lambda_n = \lambda$ independiente de n (es decir, la tasa de llegada no se afecta por el número de clientes en el sistema).

En forma similar, μ_n, la tasa media de servicio cuando hay n clientes en el sistema, también puede depender de n. En este caso, la inclusión de la dependencia con respecto a n nos permite considerar situaciones en las cuales el servidor podría acelerarse o desacelerarse dependiendo del número de clientes que esperan ser atendidos.

Si λ y μ son independientes de n, definimos la tasa de utilización del sistema, ρ, con la fórmula

$$\rho = \lambda/c\mu.$$

La tasa de utilización se interpreta como la proporción de tiempo de ocupación de cada servidor o el número esperado de clientes en servicio. Ésta debe ser menor que 1 para asegurar que el tamaño de la cola no crezca sin límite.

S2.4 FÓRMULA DE LITTLE

En esta sección mostramos algunas relaciones útiles entre los valores esperados del estado estable, L, L_q, W, y W_q. Ya que W_q es el tiempo esperado solamente en la cola, mientras W es el tiempo esperado en la cola más el tiempo esperado en el servicio, se concluye que W y W_q difieren por el tiempo esperado en el servicio. Es decir,

$$W = W_q + 1/\mu.$$

(Si la tasa media de servicio es μ, se concluye que el tiempo medio de servicio es $1/\mu$.)

La fórmula de Little se llama así en honor de John D. C. Little del MIT, quien probó que es válida bajo circunstancias muy generales. Es una relación simple pero muy útil entre los valores de L y los de W. Suponga que las tasas de llegada y de servicio son constantes; es decir, $\lambda_n = \lambda$ y $\mu_n = \mu$ para todos los valores de n. El resultado básico es

$$L = \lambda W.$$

No presentaremos una prueba formal de este resultado, sino que sólo suministraremos la siguiente explicación intuitiva. Piense en un cliente que se forma en la cola en estado estable. En el momento en que está por terminar el servicio, él[1] voltea a ver sobre su hombro a las personas que han llegado detrás suyo. En promedio habrá L clientes en el sistema. El lapso esperado que ha transcurrido desde que él se formó en la cola es, por definición, W. Ya que los clientes llegan con una tasa constante λ, se concluye que durante un tiempo W en promedio, habrá habido λW llegadas, dando $L = \lambda W$. Por ejemplo, si los clientes llegan con una tasa de 2 por minuto y cada uno ocupa un promedio de 5 minutos en el sistema, esta fórmula nos dice que en promedio habrá 10 clientes en el sistema.

Otra versión de la fórmula de Little es

$$L_q = \lambda W_q.$$

El argumento aquí es esencialmente el mismo, excepto que el cliente mira sobre su hombro al entrar en el servicio, en vez de hacerlo cuando termina el servicio.

S2.5 LAS DISTRIBUCIONES EXPONENCIALES Y DE POISSON EN LA COLA

En el capítulo 12, en el modelo de la confiabilidad, estudiamos la propiedad sin memoria de la distribución exponencial y su relación con el proceso y la distribución de Poisson. Tanto la distribución exponencial como la de Poisson juegan un papel clave en la teoría de colas, al igual que en la teoría de confiabilidad. Cuando nos referimos a llegadas puramente aleatorias en las colas, inferimos que el proceso de llegada es un proceso de Poisson. Un proceso de servicio puramente aleatorio implica que los tiempos de servicio tienen la distribución exponencial. Recomendamos que el lector repase las secciones 12.1 y 12.3, donde se discuten la distribución exponencial y el proceso de Poisson, antes de proseguir con este suplemento.

---

[1] Usamos el pronombre masculino por comodidad. Por favor lea esto y las referencias similares como él o ella.

Apartado

Resulta que el tiempo de recurrencia hacia atrás, $t - T_{i-1}$, también tiene distribución exponencial con tasa λ. El lector astuto va a intuir que algo anda mal en este punto. La suma de $t - T_{i-1}$ y $T_i - t$ da $T_i - T_{i-1}$, que es solamente un tiempo entre llegadas. Sin embargo, si $t - T_{i-1}$ es exponencial con tasa λ y $T_i - t$ también es exponencial con tasa λ, debe concluirse que $E(t - T_{i-1} + T_i - t) = E(T_i - T_{i-1}) = 2/\lambda$, ¡lo que contradice la hipótesis de que los tiempos entre llegadas son exponenciales con tasa λ! Esta contradicción evidente se conoce como la *paradoja del tiempo de espera*. Tiene que ver con el hecho de que seleccionamos un punto en el tiempo y encontramos el intervalo que incluía a ese punto en lugar de seleccionar un intervalo de manera aleatoria. En promedio, los intervalos que cubren un punto aleatorio tienen el doble de longitud. No vamos a profundizar aquí en ello, pero observe que esto mantuvo perplejos a los matemáticos por muchos años (tal como probablemente mantenga al lector). Esperamos que el lector interesado continúe investigando por su cuenta sobre este tema. Un buen punto inicial es el excelente estudio de Feller (1966), p. 11. Kleinrock (1975, p. 169) estudia la paradoja en el contexto de las colas.

Si las llegadas siguen un proceso de Poisson, esto implica que los tiempos entre llegadas tienen distribución exponencial. Debido a la propiedad sin memoria de la distribución exponencial, nos referimos al proceso de Poisson como un proceso de llegadas puramente aleatorias. Si T es una variable aleatoria que representa al tiempo entre llegadas sucesivas, entonces

$$P\{T > t\} = \exp(-\lambda t).$$

Además, el número de llegadas para cualquier instante t, digamos $A(t)$, tiene distribución de Poisson con parámetro λt.[2] Es decir,

$$P\{A(t) = n\} = \exp(-\lambda t)(\lambda t)^n/n! \qquad \text{para } n = 0, 1, 2, \ldots$$

La distribución exponencial tiene una propiedad relacionada con la propiedad sin memoria que es especialmente útil en el análisis de colas. Tiene que ver con lo que se conoce como los tiempos de recurrencia hacia adelante y hacia atrás. Sea $N(t)$ un proceso de Poisson con tasa λ, y sean $T_1, T_2, \ldots$, instantes de tiempo sucesivos entre llegadas. Considere un instante de tiempo determinista t que se sitúe entre dos tiempos sucesivos entre llegadas, por ejemplo T_{i-1} y T_i. El tiempo de recurrencia hacia adelante es la variable aleatoria $T_i - t$, o el tiempo que transcurre a partir de t hasta la siguiente llegada. La distribución exponencial es la única distribución que tiene la propiedad de que la distribución del tiempo de recurrencia hacia adelante también tenga la distribución exponencial con la tasa λ *independiente de t*. En las colas, esto implica que si un servidor está ocupado cuando llega un cliente, el lapso que transcurre hasta la terminación del servicio todavía es exponencial con tasa μ. Por ejemplo, suponga que los taxis llegan a su base de acuerdo con un proceso de Poisson. Usted llega al sitio en algún instante aleatorio y espera por el siguiente taxi. Su tiempo de espera es exponencial, con exactamente la misma distribución que el tiempo entre dos llegadas sucesivas de taxis. Esto conduce a una paradoja evidente (véase el cuadro Apartado).

Estas propiedades del proceso de Poisson y de la distribución exponencial permiten un análisis relativamente directo de los sistemas de colas. Si el proceso de llegada o el proceso de servicio no son puramente aleatorios (es decir, exponenciales), la matemática se hace mucho más compleja y los resultados disponibles se hacen más dispersos. Estudiamos lo que estos aspectos implican en la aplicación de las colas en la vida real más adelante en este suplemento.

Una pregunta importante es, ¿qué tan realistas son las hipótesis de las llegadas de Poisson y del servicio exponencial? En la práctica, cuando la población de clientes es grande y las llegadas

[2] En el capítulo 12 usamos $N(t)$ para el proceso de Poisson. Sin embargo, $N(t)$ es el estado del sistema de colas, por lo que aquí usamos $A(t)$ para el proceso de Poisson.

ocurren una a la vez, las llegadas de Poisson son una hipótesis razonable. Sin embargo, la hipótesis de tiempos de servicio con distribución exponencial es más cuestionable. En la mayoría de las circunstancias esperaríamos que entre más tiempo pase una persona en el servicio, más probable es que lo termine pronto. Vamos a estudiar las distribuciones generales de los servicios al final del suplemento.

S2.6 ANÁLISIS DE NACIMIENTO Y MUERTE PARA LA COLA M/M/1

Una notación taquigráfica de los problemas de colas tiene la forma

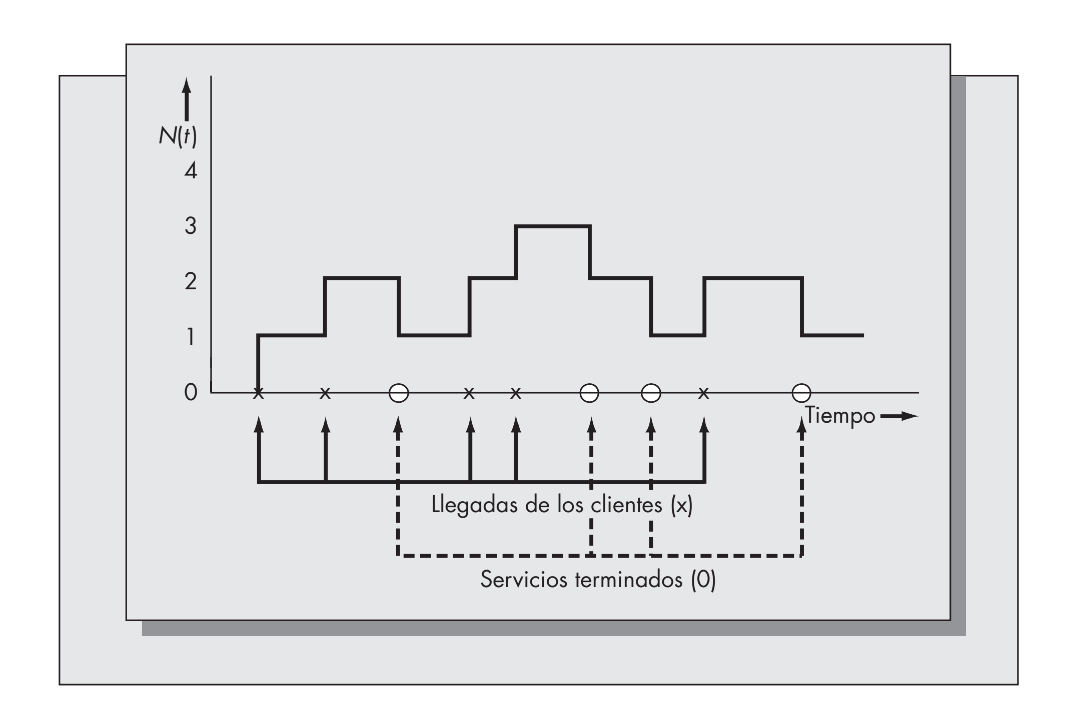
Etiqueta 1/Etiqueta 2/Número,

donde Etiqueta 1 es una abreviatura del proceso de llegada, Etiqueta 2 es una abreviatura del proceso de servicio, y Número indica el número de servidores.[3] La letra M se usa para denotar las llegadas puramente aleatorias o el servicio puramente aleatorio. Esto implica que los tiempos entre llegadas son exponenciales o los tiempos de servicio son exponenciales. La M significa markoviano, una referencia a la propiedad sin memoria de la distribución exponencial. El problema de colas más simple es el rotulado M/M/1. Otro símbolo que consideramos es G, que significa distribución general. Así G/G/s correspondería a un problema de colas en el cual la distribución de los tiempos entre llegadas fuera general, la distribución de servicio fuera general y hubiera s servidores. Hay otras etiquetas para otras distribuciones, pero no las consideramos aquí. El resto de esta sección supone una cola M/M/1.

El proceso $A(t)$, el número de llegadas hasta el instante de tiempo t, es un proceso de nacimiento puro. Se incrementa en uno para cada intervalo. El proceso $N(t)$ se conoce como un proceso de nacimiento y muerte porque aumenta y disminuye. Aumenta en uno en cada llegada y disminuye en uno en cada terminación de servicio. En la figura S2-2 se muestra una ejecución de $N(t)$. Observe que el estado del sistema aumenta en uno o disminuye en uno. La



FIGURA S2-2
El proceso $N(t)$.



[3] Existen notaciones más complejas que incluyen restricciones de la capacidad y una especificación de la disciplina de la cola. Véase, por ejemplo, Gross y Harris (1985), p. 9.

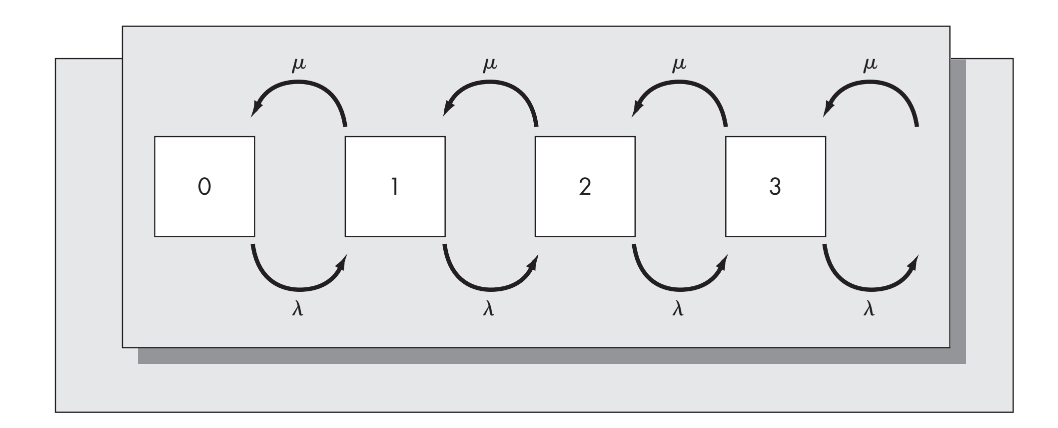


FIGURA S2-3
Cambios de estado para la cola M/M/1.

intensidad o la tasa de incremento del estado del sistema es λ y la intensidad de disminución del estado del sistema es μ.[4] Esto significa que podemos representar la tasa de cambio de estado del sistema con el diagrama de la figura S2-3.

Supongamos que el sistema ha evolucionado a una condición de estado estable. Eso significa que el estado del sistema es independiente del estado inicial. Ya que estamos en estado estable, consideramos solamente las probabilidades estacionarias P_n. La siguiente derivación se basa en el principio del balance:

> **El principio del balance:** En el estado estable, la tasa de entrada a un estado debe ser igual a la tasa de salida de un estado si existe una distribución de probabilidad del estado estable.

Considere la aplicación del principio del balance al estado 0. Entramos al estado 0 solamente a partir del estado 1. Dado que estamos en el estado 1, nos movemos del estado 1 al estado 0 con una tasa μ (véase la figura S2-3). La probabilidad de estar en el estado 1 es P_1. Se concluye que la tasa de movimiento hacia el estado 0 es μP_1. Considere la tasa de salida del estado 0. Cuando estamos en el estado 0, solamente podemos movernos al estado 1, que es lo que hacemos (cuando llega un cliente) con la tasa λ. Como la probabilidad de estar en el estado 0 es P_0, se concluye que la tasa general de salida del estado 0 es λP_0. A partir de esto obtenemos nuestra primera ecuación de balance:

$$\mu P_1 = \lambda P_0.$$

Considere el estado 1. En la figura S2-3 observamos que podemos entrar al estado 1 de dos maneras: a partir del estado 0 o a partir del estado 2. Dado que estamos en el estado 0, entramos al estado 1 con la tasa λ, y dado que estamos en el estado 2, entramos al estado 1 con la tasa μ. Se concluye que la tasa de entrada al estado 1 es $\lambda P_0 + \mu P_2$. Podemos dejar el estado 1 yendo ya sea al estado 0 si se presenta una llegada, o bien al estado 2 si se presenta la terminación del servicio. Así, la tasa de salida del estado 1 es $\lambda P_1 + \mu P_1 = (\lambda + \mu)P_1$. Se concluye que la segunda ecuación de balance es

$$\mu P_2 + \lambda P_0 = (\lambda + \mu)P_1.$$

La forma de las ecuaciones de balance *restantes* es esencialmente la misma que la de la segunda ecuación de balance. En general,

$$\mu P_{i+1} + \lambda P_{i-1} = (\lambda + \mu)P_i \qquad \text{para } 1 \le i \le \infty.$$

[4] En este punto consideramos solamente el caso en el cual las tasas de llegada y de servicio son fijas e independientes del estado del sistema. La ampliación al caso más general se considerará en la sección S2.9.

Estas ecuaciones, junto con alguna otra condición, nos permiten obtener una solución explícita de las probabilidades del estado estable. La solución consiste en expresar primero cada P_i en términos de P_0. De la primera ecuación de balance tenemos

$$P_1 = (\lambda/\mu)P_0.$$

La segunda ecuación de balance da

$$\begin{aligned}\mu P_2 &= (\lambda + \mu)P_1 - \lambda P_0 = (\lambda + \mu)(\lambda/\mu)P_0 - \lambda P_0 \\ &= (\lambda^2/\mu)P_0 + \lambda P_0 - \lambda P_0 = (\lambda^2/\mu)P_0.\end{aligned}$$

La división de ambos miembros entre μ nos da

$$P_2 = (\lambda/\mu)^2 P_0.$$

Igualmente, en general encontraremos que

$$P_i = (\lambda/\mu)^i P_0.$$

La solución se obtiene usando la condición de que

$$\sum_{i=0}^{\infty} P_i = 1,$$

ya que $P_0, P_1, P_2, \ldots$ constituye una distribución de probabilidad para los estados del sistema.

Con la sustitución de cada P_i, tenemos

$$\sum_{i=0}^{\infty} (\lambda/\mu)^i P_0 = 1.$$

Sea $\rho = \lambda/\mu$ la tasa de utilización. Para que exista una solución, debe ser verdad que $\rho < 1$. En ese caso

$$\sum_{i=0}^{\infty} \rho^i = 1/(1 - \rho),$$

que se conoce como la serie geométrica, de la cual obtenemos

$$P_0 = (1 - \rho)$$

y

$$P_i = \rho^i(1 - \rho) \qquad \text{para } i = 1, 2, 3, \ldots$$

(Esta fórmula también es válida si $i = 0$.)

Esta distribución, conocida como distribución geométrica, se ilustra en la figura S2-4. Hay varios aspectos de este resultado que son interesantes y sorprendentes. Primero, la distribución geométrica es la analogía discreta de la distribución exponencial. (La exponencial es la distribución tanto de los tiempos entre llegadas como de los tiempos de servicio.) Segundo, la probabilidad del estado i es una función decreciente de i, como se ilustra en la figura S2-4, siempre que $\rho < 1$. A medida que ρ se aproxima a uno, la varianza aumenta y la distribución se "dilata" (los valores grandes se hacen más probables), y a medida que ρ se aproxima a cero las probabilidades asociadas con los valores más grandes tienden a cero con mayor rapidez. Esto significa que el estado más probable *siempre* es el estado 0 (¡siempre que $\rho < 1$!). Éste es un resultado muy sorprendente. A medida que ρ se aproxima a uno, las colas se hacen cada vez más largas. Uno pensaría que la probabilidad de n en el sistema para algún valor grande de n sería mayor que la probabilidad de cero en el sistema cuando ρ esté cerca de uno. Éste no es el caso. La verdad es que para ρ cercano a uno, la probabilidad de que el sistema esté en el estado cero se acerca a la probabilidad de que el sistema esté en el estado 1 o bien, en el estado 2, por ejemplo, mientras que para ρ cercano a cero la probabilidad de que el sistema

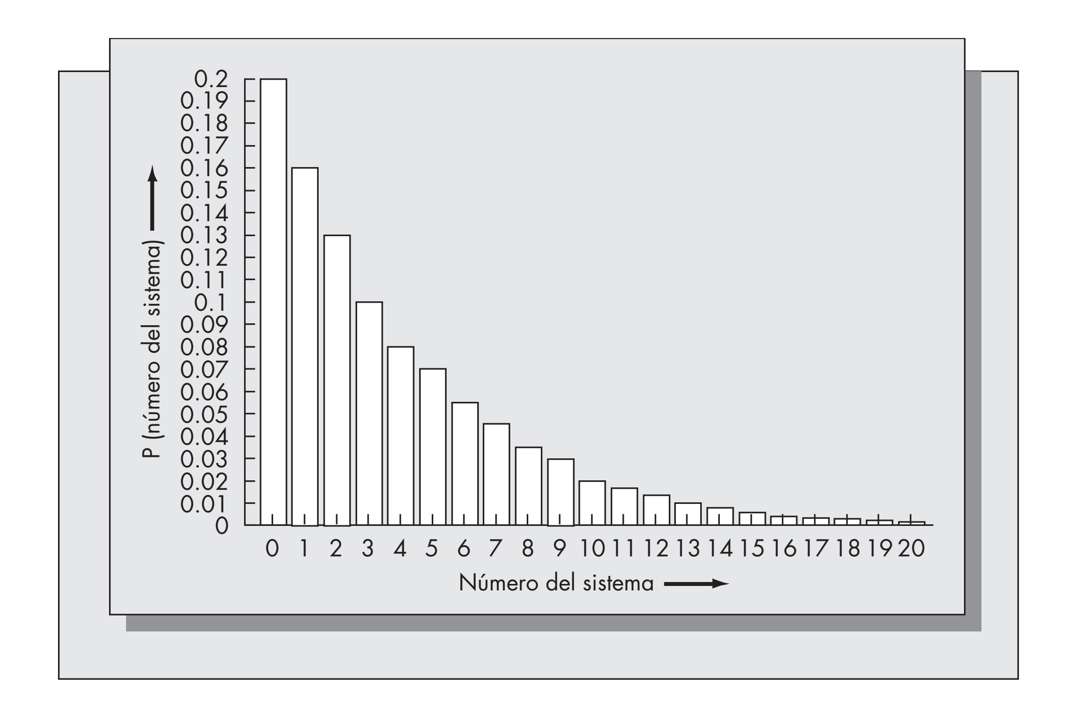


FIGURA S2-4
La distribución geométrica del número del sistema para una cola M/M/1 ($\rho = .8$).

esté en el estado cero es mucho mayor que la probabilidad de que esté en el estado 1 o 2. No obstante, este fenómeno es válido *solamente* para servicios y llegadas exponenciales.

S2.7 CÁLCULO DE LAS MEDIDAS ESPERADAS DE SISTEMA PARA LA COLA M/M/1

En esta sección calculamos L, L_q, W, y W_q. El valor esperado de una variable aleatoria es la suma de sus valores ponderados por sus probabilidades respectivas. Se concluye que el número promedio de clientes en el sistema en estado estable, L, es

$$L = \sum_{i=0}^{\infty} iP_i = \sum_{i=0}^{\infty} i(1-\rho)\rho^i = (1-\rho)p \sum_{i=0}^{\infty} i\rho^{i-1}.$$

Para terminar el cálculo, usamos el hecho de que

$$\sum_{i=0}^{\infty} i\rho^{i-1} = \frac{d}{}\left(\sum_{i=0}^{\infty} \rho^i\right) = \frac{d}{}\left(\frac{1}{1-\rho}\right) = \frac{1}{(1-\rho)^2}.$$

Se concluye que

$$L = \frac{(1-\rho)\rho}{(1-\rho)^2} = \frac{\rho}{1-\rho}.$$

Para el caso de L_q, observamos que el número en la cola es exactamente uno menos que el número en el sistema siempre que haya cuando menos uno en el sistema. Se concluye que

$$L_q = \sum_{i=1}^{\infty}(i-1)P_i = \sum_{i=1}^{\infty} iP_i - \sum_{i=1}^{\infty} P_i$$
$$= L - (1 - P_0) = L - \rho = \rho^2/(1-\rho).$$

Si se conocen L y L_q, podemos obtener W y W_q directamente de la fórmula de Little. De la fórmula de Little $W = L/\lambda$, dando

$$W = \frac{\rho}{\lambda(1-\rho)}.$$

De forma similar, $W_q = L_q/\lambda$, que da

$$W_q = \frac{\rho^2}{\lambda(1-\rho)}.$$

S2.8 LA DISTRIBUCIÓN DEL TIEMPO DE ESPERA

En esta sección, derivamos la distribución del tiempo de espera W para un cliente aleatorio que se incorpora a la cola en estado estable. Por comodidad, en esta sección usamos el símbolo W para referir la variable aleatoria, en vez de su esperanza. La interpretación apropiada de W quedará clara en el contexto.

Para derivar la distribución de W condicionamos el número de clientes en el sistema para el estado estable, n, y quitamos la condición al multiplicar por la probabilidad P_n. Suponga que un cliente que se incorpora a la cola en un punto aleatorio en el tiempo encuentra que n clientes ya están en el sistema. Así, ese cliente debe esperar a que se terminen n servicios antes de ser atendido. Como W es el tiempo total en el sistema, significa que en este caso W será la suma de $n + 1$ terminaciones de servicio. Sean $S_1, S_2, \ldots$ los tiempos de los servicios sucesivos. Por suposición, estas variables aleatorias son mutuamente independientes y tienen distribución exponencial con media común μ. El tiempo para $n + 1$ terminaciones de servicio es $S_1 + S_2 + \cdots + S_{n+1}$, que sabemos tiene la distribución Erlang con parámetros μ y $n + 1$ (véase la sección 12.3).

Esto es,

$$P\{W > t \,|\, n \text{ en el sistema}\} = \sum_{k=0}^{n} \frac{e^{-\mu t}(\mu t)^k}{k!}.$$

Gracias a la sección anterior sabemos que la probabilidad condicional de n en el sistema en el estado estable, P_n, tiene distribución geométrica. Al sustituir $\rho = \lambda/\mu$, podemos escribir P_n en la forma

$$P_n = \left(\frac{\mu-\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n.$$

Quitar la condición de P_n nos da

$$\begin{aligned} P\{W > t\} &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{e^{-\mu t}(\mu t)^k}{k!}\left(\frac{\mu-\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n \\ &= \sum_{k=0}^{\infty}\sum_{n=k}^{\infty} \frac{e^{-\mu t}(\mu t)^k}{k!}\left(\frac{\mu-\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n \\ &= \frac{\mu-\lambda}{\mu}e^{-\mu t}\sum_{k=0}^{\infty}\frac{(\mu t)^k}{k!}\sum_{n=k}^{\infty}\left(\frac{\lambda}{\mu}\right)^n. \end{aligned}$$

Usando el hecho de que

$$\sum_{n=k}^{\infty}\left(\frac{\lambda}{\mu}\right)^n = \left(\frac{\lambda}{\mu}\right)^k \frac{1}{1-\lambda/\mu},$$

y sustituyendo esto en la ecuación anterior, después de simplificar, nos da

$$P\{W > t\} = e^{-\mu t}\sum_{k=0}^{\infty}\frac{(\lambda t)^k}{k!} = e^{-\mu t}e^{+\lambda t} = e^{-(\mu-\lambda)t}.$$

El término de la sumatoria es igual a $e^{+\lambda t}$ porque es la expansión en serie de Taylor de e (y porque las probabilidades de Poisson suman uno). Lo que hemos mostrado es el resultado sorprendente de que W tiene distribución exponencial negativa con parámetro $\mu - \lambda$. Esto implica que W tiene la propiedad sin memoria. Es decir, suponga que un cliente ya ha estado esperando durante s unidades de tiempo. La probabilidad de que deba esperar cuando menos t unidades adicionales de tiempo es la misma que la probabilidad de que un cliente que se acaba de incorporar espere cuando menos t unidades de tiempo. ¡Este resultado no es intuitivo y es más bien deprimente para el pobre cliente, quien ya ocupó una gran cantidad de tiempo esperando el servicio!

No vamos a presentar la derivación (es similar a la antes dada), pero afirmaremos que la distribución de W_q es esencialmente exponencial con la función de distribución acumulada complementaria

$$P\{W_q > t\} = \rho e^{-(\mu-\lambda)t} \qquad \text{para todo } t \geq 0.$$

Observe que la probabilidad de que el tiempo de espera en la cola sea cero es positiva. Equivale a la probabilidad de que el sistema esté vacío, P_0. Es decir,

$$P\{W_q = 0\} = P_0 = 1 - \rho.$$

Ejemplo *S2.1*

Los clientes llegan uno a la vez, completamente en forma aleatoria, a un cajero automático (ATM) cuya tasa es de seis clientes por hora. Los clientes toman un promedio de cuatro minutos para terminar sus transacciones, y los datos históricos han mostrado que los tiempos de servicio se ajustan a la distribución exponencial negativa. Los clientes hacen cola sobre la base de primero en llegar, primero en salir. Suponga que solamente hay un ATM.

a) Encuentre las siguientes medidas esperadas del desempeño para este sistema: el número esperado de clientes en el sistema, el número esperado de clientes que esperan el servicio, el tiempo esperado en el sistema y el tiempo esperado en la cola.

b) ¿Cuál es la probabilidad de que haya más de cinco personas en el sistema para un punto aleatorio en el tiempo?

c) ¿Cuál es la probabilidad de que el tiempo de espera en la cola sobrepase los 10 minutos?

d) Dados estos resultados, ¿piensa que la administración deba considerar anexar otro ATM?

Solución

La afirmación de que los clientes llegan uno a la vez en tiempos completamente aleatorios implica que el proceso de entrada es un proceso de Poisson. La tasa de llegada es $\lambda = 6$ por hora. El tiempo medio de servicio es de 4 minutos $= 1/15$ hora, por lo que la tasa de servicio es $\mu = 15$ por hora. La tasa de utilización es $\rho = \lambda/\mu = 6/15 = 2/5 = 0.4$.

a) $L = \rho/(1 - \rho) = (2/5)/(3/5) = 2/3$ $(= 0.6667)$.
$L_q = \rho L = (2/5)(2/3) = 4/15$ $(= 0.2667)$.
$W = L/\lambda = (2/3)/6 = 2/18 = 1/9$ hora (6.6667 minutos).
$W_q = L_q/\lambda = (4/15)/6 = 4/90 = 2/45$ hora (2.6667 minutos).

b) Aquí estamos interesados en $P\{L > 5\}$.
En general,

$$P\{L > k\} = \sum_{n=k+1}^{\infty} P_n = \sum_{n=k+1}^{\infty} (1-\rho)\rho^n = (1-\rho)\sum_{n=k+1}^{\infty}\rho^n$$
$$= (1-\rho)\rho^{k+1}(1/(1-\rho)) = \rho^{k+1}.$$

Entonces, $P\{L > 5\} = \rho^6 = (0.4)^6 = 0.0041$.

c) Aquí estamos interesados en $P\{W_1 > 1/6\}$.

$$P\{W_q > t\} = \rho e^{-(\mu-\lambda)t} = 0.4e^{-(15-6)(1/6)} = 0.4e^{-1.5} = 0.0892.$$

d) La respuesta no es obvia. Si examinamos las medidas esperadas del desempeño, parecería que el servicio suministrado es razonable. El número esperado de clientes en el sistema es menor que 1 y el tiempo de espera promedio en la cola es menor que 3 minutos. Sin embargo, en el inciso *c* vemos que la proporción de clientes que debe esperar más de 10 minutos por servicio es casi del 10 por ciento. Ello significa que probablemente haya muchos clientes enojados, aun cuando en promedio el sistema parezca bueno. Esto ilustra la falla que reside en considerar sólo valores esperados cuando se evalúan los sistemas de servicio de colas.

S2.9 SOLUCIÓN DEL CASO GENERAL

En esta sección consideramos una solución de la cola general markoviana; es decir, cuando los tiempos entre llegadas sucesivas tienen distribución exponencial y los tiempos de servicio también la tienen. Sin embargo, dejaremos que las tasas de llegada y las de servicio dependan del estado. Varias versiones del modelo M/M/1 son casos especiales de éste.

El diagrama de transición es el mismo que el que se ilustra en la figura S2-3, excepto que tanto μ como σ son dependientes del estado (véase la figura S2-5). El principio de la ecuación de balance aplicado a este sistema arroja

$$\mu_1 P_1 = \lambda_0 P_0,$$
$$\lambda_0 P_0 + \mu_2 P_2 = (\lambda_1 + \mu_1)P_1,$$
$$\lambda_1 P_1 + \mu_3 P_3 = (\lambda_2 + \mu_2)P_2,$$

etcétera.

Expresando cada una de las probabilidades de estado en términos de P_0 como lo hicimos anteriormente conduce a lo siguiente:

$$P_1 = \frac{\lambda_0}{\mu_1}P_0,$$

$$P_2 = \frac{\lambda_1\lambda_0}{\mu_2\mu_1}P_0,$$

$$P_3 = \frac{\lambda_2\lambda_1\lambda_0}{\mu_3\mu_2\mu_1}P_0,$$

y así sucesivamente.

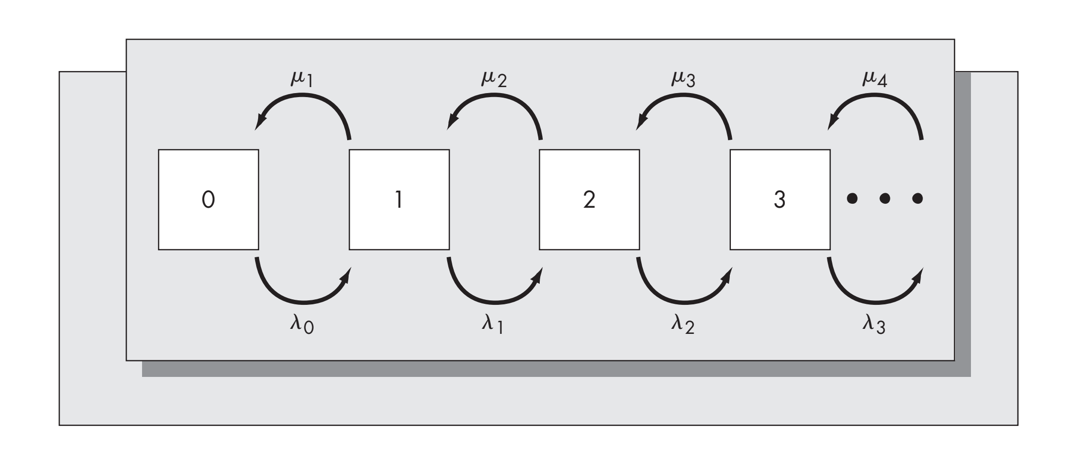


FIGURA S2-5
Cambios de estado para la cola M/M/1 con tasas de servicio y de llegada dependientes del estado.

Defina

$$A_n = \frac{\lambda_{n-1}\lambda_{n-2} \dots \lambda_0}{\mu_n\mu_{n-1} \dots \mu_1}$$

de modo que

$$P_n = A_n P_0 \qquad \text{para } n = 1, 2, 3, \dots$$

Nuevamente usando el hecho de que $P_0, P_1, \dots$ es una distribución de probabilidad, tenemos que

$$\sum_{n=0}^{\infty} P_n = 1.$$

Esto se traduce a la condición de definición de P_0 como

$$P_0 = \frac{1}{1 + \sum_{n=1}^{\infty} A_n}.$$

Las diferentes medidas de servicio pueden obtenerse aplicando sus definiciones. En especial, L, el número esperado en el sistema, está dado por

$$L = \sum_{n=0}^{\infty} nP_n.$$

Si se supone que hay c servidores, entonces L_q, el número esperado en la cola, está dado por

$$L_q = \sum_{n=c}^{\infty} (n - c)P_n.$$

La fórmula de Little todavía es aplicable y puede usarse para encontrar los tiempos de espera previstos dado el número esperado en el sistema y el número esperado en la cola. No obstante, para aplicar la fórmula de Little cuando la tasa de llegada depende del estado, debemos determinar el promedio general esperado de la *tasa* de llegada, o la tasa de llegada efectiva, que llamamos λ_{eff}. Ya que la tasa de llegada es λ_n cuando el sistema está en el estado n, se concluye que la tasa de llegada efectiva es

$$\lambda_{eff} = \sum_{n=0}^{\infty} \lambda_n P_n.$$

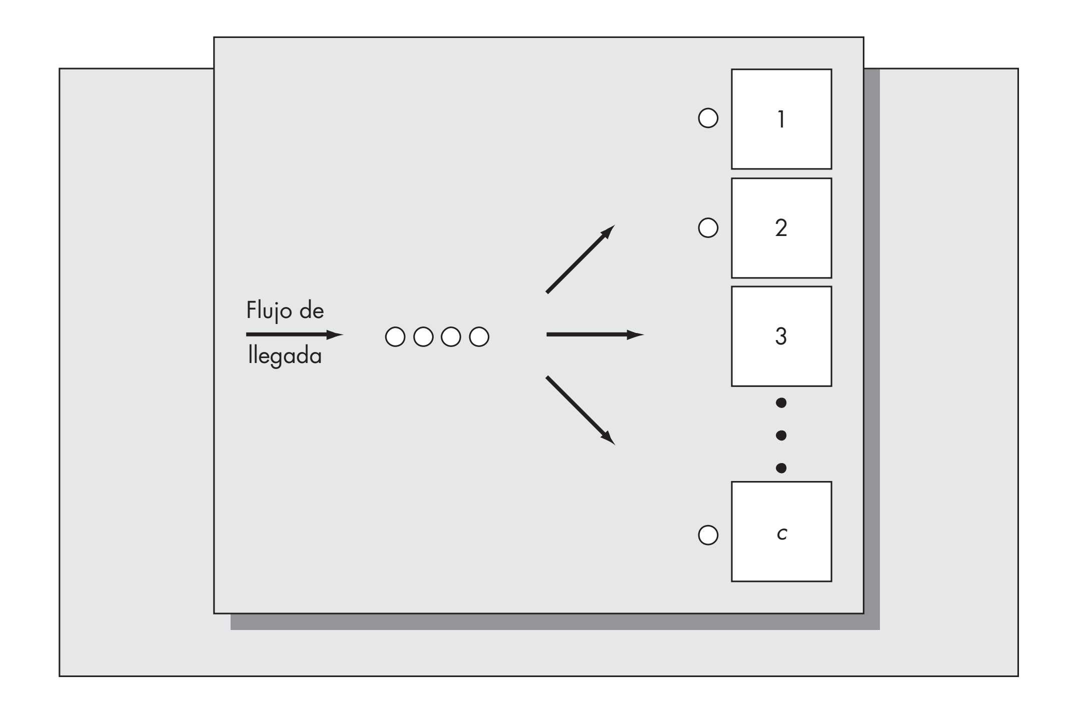
S2.10 SERVIDORES MÚLTIPLES EN PARALELO: LA COLA M/M/*c*

En esta sección, consideramos cómo pueden usarse las ecuaciones de la sección S2.9 para resolver varias configuraciones de la cola con tiempos exponenciales entre llegadas y tiempos de servicio exponencial.

Primero, considere la cola M/M/c; es decir, el caso en el cual hay c servidores en paralelo. Este caso se ilustra en la figura S2-6. Cuando llegan los clientes, hacen cola en una sola fila. El siguiente cliente en la fila es atendido por el siguiente servidor disponible. Resulta que éste es solamente un caso especial del modelo dependiente del estado que se derivó anteriormente, pero para demostrar eso, necesitamos establecer el siguiente resultado.

Suponga que m servidores están ocupados en un punto aleatorio en el tiempo y también suponga que el tiempo de servicio de cada uno de los servidores es exponencial con tasa μ. La pregunta es, ¿cuál es la distribución del tiempo hasta que termina el siguiente servicio? Sean $T_1, T_2, \dots, T_m$ los tiempos de servicio de los clientes que actualmente están en servicio. Por hipótesis, éstas son variables aleatorias exponenciales independientes. Aún más, si t es un punto aleatorio en el tiempo, el tiempo restante en servicio desde t hasta el final de la terminación

FIGURA S2-6
c servidores
en paralelo.



del servicio para cada uno de los clientes también es exponencial con la misma distribución. (Ésta es una consecuencia de las propiedades de la distribución exponencial estudiada anteriormente.)

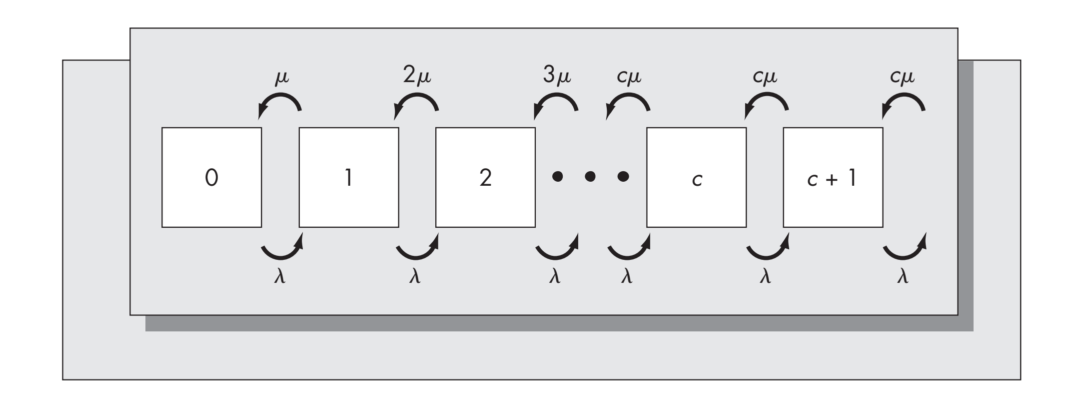
Se concluye que el tiempo hasta la siguiente terminación del servicio, digamos T, se distribuye como el *mínimo* de $T_1, T_2, \ldots, T_m$. El resultado que necesitamos analizar en este caso es

> **Resultado:** Sean $T_1, T_2, \ldots, T_m$ variables aleatorias exponenciales independientes con una distribución exponencial común con tasa μ, y definamos $T = \text{mín}(T_1, T_2, \ldots, T_n)$. Entonces T también tiene distribución exponencial, con tasa $m\mu$.

(Este resultado se prueba en el capítulo 12, en el contexto de los sistemas de componentes en serie sujetos a una falla exponencial.)

¿Por qué es importante este resultado? Significa que la distribución entre las salidas de los clientes todavía es exponencial y que pueden aplicarse los métodos usados en la sección S2.9. Si hay c servidores, entonces se concluye que el diagrama de la tasa de transición es como se ilustra en la figura S2-7.

FIGURA S2-7
Diagrama de tasa
de transición cuando
existen c servidores
en paralelo.



Al comparar las figuras S2-5 y S2-7, vemos que

$$\begin{aligned}\mu_1 &= \mu\\ \mu_2 &= 2\mu\\ &\vdots\\ \mu_c &= c\mu\\ \mu_{c+1} &= c\mu\\ \mu_{c+2} &= c\mu\end{aligned}$$

y $\lambda_i = \lambda$ para todo $i = 0, 1, 2, \ldots$

Al sustituir, se concluye que

$$\begin{aligned}P_1 &= \frac{\lambda}{\mu} P_0,\\ P_2 &= \frac{1}{2}\left(\frac{\lambda}{\mu}\right)^2 P_0,\\ P_3 &= \frac{1}{(3)(2)}\left(\frac{\lambda}{\mu}\right)^3 P_0,\\ &\vdots\\ P_n &= \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n P_0, \qquad \text{para } 0 \le n \le c.\end{aligned}$$

Cuando $n > c$ obtenemos

$$P_n = \frac{1}{c!c^{n-c}}\left(\frac{\lambda}{\mu}\right)^n P_0, \qquad \text{para } n > c.$$

La sustitución de estas probabilidades de estado da lo siguiente para P_0:

$$P_0 = \left\{\sum_{n=0}^{c-1} \frac{1}{n!}\frac{\lambda^n}{\mu} + \sum_{n=c}^{\infty} \frac{1}{c!c^{n-c}}\left(\frac{\lambda}{\mu}\right)^n\right\}^{-1}.$$

Esto puede simplificarse observando que el segundo término es una serie geométrica. Estableciendo $\rho = \lambda/c\mu$, se puede mostrar después de un poco de manipulación algebraica que

$$P_0 = \left\{\sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!}(1-\rho)^{-1}\right\}^{-1}.$$

Al calcular las medidas de desempeño estándar, resulta que L_q tiene la forma más simple. Nuevamente no vamos a presentar los detalles, sino sólo los resultados. Las derivaciones son similares a las presentadas en secciones anteriores.

$$\begin{aligned}L_q &= \frac{c^c\rho^{c+1}}{c!(1-\rho)^2} P_0.\\ L &= L_q + c\rho.\\ W_q &= L_q/\lambda.\\ W &= W_q + 1/\mu.\end{aligned}$$

Al igual que con la cola de un solo servidor, se requiere la condición de que $\rho < 1$ para garantizar que la cola no crezca sin límite.

Ejemplo *S2.2*

La peluquería Tony es atendida y administrada por su dueño, Anthony Jones, quien ha cortado el cabello por más de 20 años. Anthony no hace citas, de modo que el patrón de llegada de los clientes es en esencia aleatorio. Tradicionalmente, la tasa de llegada había sido aproximadamente un cliente cada 50 minutos. Hace dos meses, el periódico local publicó un artículo acerca de Anthony, que mejoró los negocios sustancialmente. En la actualidad, la tasa de llegada es más cercana a un cliente cada 35 minutos. Los cortes de pelo requieren un promedio de 25 minutos, pero los tiempos varían considerablemente dependiendo de las necesidades del cliente. Un corte puede requerir tan poco como cinco minutos, pero un champú y un peinado pueden tomar tanto como una hora o más. Por esta razón, la distribución exponencial parece suministrar un ajuste razonablemente bueno de la distribución del tiempo de servicio.

Los clientes de Anthony siempre han sido pacientes, pero desde que el negocio floreció, algunos se han quejado de que las esperas son demasiado largas. Anthony está considerando incorporar a su primo Marvin en el negocio para mejorar el servicio al cliente. Suponga que Marvin corta el pelo con la misma rapidez que Anthony.

a) ¿Qué tanto ha declinado la calidad del servicio desde que más clientes han comenzado a ir a la peluquería de Anthony?

b) ¿Qué tanta mejoría en el desempeño del sistema es probable que vean los clientes con un peluquero adicional en el salón?

Solución

Primero vamos a determinar las diferentes medidas de desempeño del sistema antes de la aparición del artículo del periódico. El tiempo promedio entre llegadas fue uno cada 50 minutos, lo que da una tasa de llegada de

$$\lambda = 60/50 = 1.2 \text{ llegadas por hora.}$$

Cada corte de pelo requiere un promedio de 25 minutos, lo que se traduce en una tasa de servicio de

$$\mu = 60/25 = 2.4 \text{ cortes de pelo por hora.}$$

Se concluye que $\rho = \lambda/\mu = 1.2/2.4 = 0.5$. (Es decir, Tony estaba ocupado la mitad del tiempo.) Los valores de las medidas de desempeño son

$$L = \rho/(1 - \rho) = 0.5/0.5 = 1.$$

$$L_q = \rho L = 0.5.$$

$$W = L/\lambda = 1/1.2 = 0.8333 \text{ hora.}$$

$$W_q = L_q/\lambda = 0.5/1.2 = 0.4167 \text{ hora.}$$

Esto significa que, en promedio, los clientes originales esperaron (0.4167)(60) = 25 minutos por un corte de pelo.

Después que apareció el artículo, la tasa de llegada aumentó a un cliente cada 35 minutos. Esto implica que λ cambió a $60/35 = 1.7143$ y $\rho = 0.7143$. Ahora las medidas de desempeño son

$$L = 0.7143/(1 - 0.7143) = 2.5.$$

$$L_q = \rho L = (0.7143)(2.5) = 1.7857.$$

$$W = L/\lambda = 2.5/1.7143 = 1.458 \text{ horas.}$$

$$W_q = L_q/\lambda = 1.78/1.7143 = 1.0383 \text{ horas.}$$

Es evidente que los clientes tienen una queja válida. Un cliente tiene que esperar en promedio más de una hora antes de cortarse el pelo. De hecho, ya que la distribución de W_q es exponencial, muchos tendrían que esperar bastante más que esto.

b. Agregar un peluquero mejora enormemente el desempeño del sistema. Con dos peluqueros, tenemos

$$\rho = \lambda/(c\mu) = 1.7143/(2)(2.4) = 0.3571. \qquad (c\rho = 0.7143).$$

$$P_0 = \left\{1 + 0.7143 + \frac{(0.7143)^2}{2!} \; \frac{1}{1 - 0.3571}\right\}^{-1}$$

$$= (2.111)^{-1} = 0.4737.$$

Se concluye que

$$L_q = \frac{(2)^2(0.3571)^3}{2!(1 - 0.3571)^2}(0.4737) = 0.0522.$$

$$L = L_q + c\rho = 0.0522 + 0.7143 = 0.7665.$$

$$W_q = L_q/\lambda = 0.0522/1.7143 = 0.0304 \text{ hora (1.82 minutos).}$$

$$W = W_q + 1/\mu = 0.0304 + 0.4167 = 0.4471 \text{ hora (por 27 minutos).}$$

Con un solo peluquero, los clientes podrían suponer que deben esperar más de una hora por un corte de pelo. Con la adición de otro peluquero, ¡el tiempo se reduce a menos de dos minutos en promedio!

S2.11 LA COLA M/M/1 CON UNA CAPACIDAD FINITA

Otra versión especial de la cola general M/M/1, con un servicio y tasas de llegada dependientes del estado, es el caso en el cual hay un área de espera finita. Si se presentan llegadas cuando el área de espera está llena, éstas se rechazan. Los problemas de este tipo son comunes en los sistemas de servicios tales como los restaurantes, los cines y las salas de concierto. También pueden presentarse en los sistemas de manufactura en los cuales los amortiguadores entre los centros de trabajo tienen una capacidad finita. Éste es el caso, por ejemplo, con los sistemas JIT. (Véase el estudio de JIT en el capítulo 7.)

Suponga que el número máximo de clientes que se permiten en el sistema es K. El diagrama de la tasa de transición para este caso es exactamente el mismo que el que se ilustra en la figura S2-3 salvo que las transiciones no ocurren más allá del estado K. Ya que el diagrama de la tasa de transición es el mismo hasta el estado K, las ecuaciones de balance van a arrojar la misma relación entre P_n y P_0 para $n = 1, 2, \ldots, K$. Es decir,

$$P_n = \rho^n P_0 \qquad \text{para } n = 1, 2, 3, \ldots, K.$$

P_0 se encuentra a partir de

$$\sum_{n=0}^{K} P_n = 1,$$

lo que da

$$P_0 = \left(\sum_{n=0}^{K} \rho^n\right)^{-1}.$$

Una clara expresión para la serie geométrica finita se obtiene de la siguiente manera:

$$\sum_{n=0}^{K} \rho^n = \sum_{n=0}^{\infty} \rho^n - \sum_{n=k+1}^{\infty} \rho^n = \frac{1}{(1-\rho)} - \frac{\rho^K}{(1-\rho)}$$

$$= \frac{1-\rho^K}{1-\rho}.$$

Se concluye que

$$P_0 = \frac{1-\rho}{1-\rho^{K+1}},$$

de donde obtenemos

$$P_n = \frac{(1-\rho)\rho^n}{1-\rho^{K+1}} \qquad \text{para } n = 0, 1, 2, \ldots, K.$$

Para el caso de la sala de espera finita, no es necesario que $\rho < 1$. De hecho, P_n adopta este valor para *todos* los valores de $\rho \neq 1$. Si $\rho = 1$, resulta que todos los estados son igualmente probables, de modo que

$$P_n = 1/(K+1) \qquad \text{para } 0 \leq n \leq K \qquad \text{(solamente cuando } \rho = 1).$$

La fórmula de Little todavía puede aplicarse, pero debemos usar un valor modificado para la tasa de llegada porque no se permite la entrada al sistema a todos los clientes que lleguen. Cuando hay K o más en el sistema, la tasa de llegada es cero, de modo que la tasa de llegada general es menor que λ. La tasa de llegada efectiva, λ_{eff}, se calcula como sigue:

$$\begin{aligned}\lambda_{eff} &= \lambda P\{\text{Número en el sistema} < K\} + 0P\{\text{Número en el sistema} = K\}\\ &= \lambda(1 - P\{\text{número en el sistema} = K\})\\ &= \lambda(1 - P_K).\end{aligned}$$

Las medidas de desempeño se obtienen de L, el número esperado en el sistema en estado estable. L se encuentra a partir de

$$\begin{aligned}L &= \sum_{n=0}^{K} nP_n\\ &= \sum_{n=0}^{K} \frac{(1-\rho)}{1-\rho^{K+1}}\, n\rho^n.\end{aligned}$$

El cálculo prosigue al observar que

$$\sum_{n=0}^{K} n\rho^{n-1} = \frac{d}{} \sum_{n=0}^{K} \rho^n.$$

Usando la expresión anterior para la suma geométrica finita, finalmente obtenemos

$$L = \frac{\rho}{1-\rho} - \frac{(K+1)\rho^{K+1}}{1-\rho^{K+1}}.$$

Las medidas restantes de desempeño se encuentran a partir de

$$L_q = L - (1 - P_0).$$
$$W = L/\lambda_{eff}.$$
$$W_q = L_q/\lambda_{eff}.$$

Pueden obtenerse fórmulas similares para el caso de una cola de capacidad finita y para servidores paralelos múltiples. (Véase, por ejemplo, Hillier y Lieberman, 1990.)

Ejemplo *S2.3*

Una atracción popular en la costa de Nueva Jersey es un artista callejero que dibuja una caricatura en aproximadamente cinco minutos. Sin embargo, ya que los tiempos requeridos para cada dibujo varían mucho, la distribución exponencial los describe con exactitud. Las personas están deseosas de esperar su turno, pero cuando hay más de 10 esperando por una caricatura, los clientes son rechazados y se les pide que vuelvan más tarde. En los tiempos pico pueden esperarse tanto como 20 clientes por hora. Suponga que los clientes acuden de forma completamente aleatoria para la tasa de llegada pico.

a) ¿Qué proporción del tiempo se encuentra la cola a su máxima capacidad?

b) ¿Cuántos clientes se rechazan en promedio? Determine las medidas de desempeño para este sistema de colas.

c) Si el área de espera doblara su tamaño, ¿cómo afectaría eso a sus respuestas de los incisos *a* y *b*?

Solución

La tasa de llegada es $\lambda = 20$ por hora y la tasa de servicio es $\mu = 12$ por hora, de modo que $\rho = 20/12 = 1.667$. El número máximo en el sistema es $K = 11$ (10 en la cola más el cliente que está siendo atendido).

a) La probabilidad de que el sistema esté lleno es PK, que aparece dada por

$$P_K = \frac{(1-\rho)\rho^K}{1-\rho^{K+1}} = \frac{(1-1.667)(1.667)^{11}}{1-1.667^{12}} = \frac{-184.25}{-459.5} = 0.40.$$

b) Como la tasa de llegada es 20 por hora y el sistema está lleno 40 por ciento del tiempo, durante los periodos pico se rechazan (20)(0.40) = 8 clientes por hora. Esto da $\lambda_{eff} = 12$ por hora.

$$L = \frac{\rho}{1-\rho} - \frac{(K+1)\rho^{K+1}}{1-\rho^{K+1}} = \frac{1.667}{1-1.667} - \frac{(12)(1.667)^{12}}{1-1.667^{12}}$$
$$= -2.5 - (-12.03)$$
$$= 9.53.$$

Necesitamos determinar P_0 para calcular L_q.

$$P_0 = 1 - \frac{1-\rho}{1-\rho^{K+1}} = \frac{1-1.667}{1-1.667^{12}} = 0.00145,$$

lo que da $L_q = L - (1 - P_0) = 9.53 - (1 - 0.00145) = 8.53$.

Observe que el valor pequeño de P_0 implica que el sistema rara vez está vacío. En especial, ¡el artista está desocupado solamente 1.45 por ciento del tiempo!

Anteriormente mostramos que $\lambda_{eff} = 12$, de modo que

$$W = L/\lambda_{eff} = 9.53/12 = 0.7942 \text{ hora (cerca de 48 minutos).}$$
$$W_q = L_q/\lambda_{eff} = 8.53/12 = 0.7108 \text{ hora (cerca de 43 minutos).}$$

c) Si el área de espera doblara su tamaño, entonces $K = 21$. En ese caso P_K está dado por

$$P_{21} = \frac{(1 - 1.667)(1.667)^{21}}{1 - 1.667^{22}} = \frac{-30{,}533.28}{-76{,}309.3} = 0.40.$$

Es interesante que duplicar la capacidad de la cola no afecta la probabilidad de que el sistema esté lleno. La razón es que, como la tasa de llegada sobrepasa a la tasa de servicio, el sistema alcanza su capacidad rápidamente en los dos casos. En ambos casos la tasa efectiva de llegada λ_{eff} es aproximadamente igual a la tasa de servicio (aunque siempre será verdad que $\lambda_{eff} < \mu$). Aun para valores mucho mayores que K, P_K es 0.4 en este ejemplo.

S2.12 RESULTADOS PARA SERVICIOS DE DISTRIBUCIÓN NO EXPONENCIALES

Una hipótesis que se requiere para todos los modelos estudiados hasta ahora es que la distribución del tiempo de servicio es exponencial. En muchos casos, no se garantiza que esta hipótesis se cumpla. Uno esperaría que los tiempos de servicio rara vez fueran exponenciales porque la distribución exponencial tiene la propiedad sin memoria: el lapso restante en servicio tendría que ser independiente del tiempo ya transcurrido. Se pensaría que una distribución modal, tal como la normal o la de Erlang, sería un modelo más exacto de los tiempos de servicio en la mayoría de las circunstancias.

Por esa razón, nos interesa mucho el modelo M/G/1. Aquí la G representa general, lo que implica *cualquier* distribución del tiempo de servicio. Como perdemos la propiedad sin memoria cuando los servicios no son exponenciales, el método del diagrama de tasas ya no se puede aplicar. Las matemáticas se hacen un poco más complejas (el lector que tenga interés debe consultar Gross y Harris, 1985, por ejemplo, para una presentación manejable de este modelo).

Las medidas de desempeño se derivan de la fórmula de Pollazek-Khintchine (*P-K*), que da el número esperado en la cola en términos de la media y la varianza de la distribución de los tiempos de servicio. Sea $E(T)$ el tiempo de servicio esperado y $\mathrm{Var}(T)$ la varianza del tiempo de servicio. Entonces la fórmula *P-K* para el número esperado en la cola es

$$L_q = \frac{\lambda^2\{\mathrm{Var}(T) + E(T)^2\}}{2(1 - \lambda E(T))}.$$

Usando nuestra notación anterior, $E(T) = 1/\mu$ y $\lambda E(T) = \lambda/\mu = \rho$, y haciendo $\mathrm{Var}(T) = \sigma^2$, esta fórmula se transforma en

$$L_q = \frac{\lambda^2\sigma^2 + \rho^2}{2(1 - \rho)}.$$

Observe que L_q (y por tanto L, W y W_q) aumenta en σ^2. Esto implica que la congestión aumenta a medida que aumenta la varianza del tiempo de servicio. Así, las medidas de desempeño para un sistema pueden mejorar al disminuir la varianza del tiempo de servicio, aun cuando el tiempo medio de servicio no cambie. (Si los tiempos de servicio tienen distribución exponencial, $\sigma^2 = 1/\mu^2$ y la fórmula para L_q se reduce a la obtenida para la cola M/M/1.)

Ilustramos la aplicación de la fórmula *P-K*.

Ejemplo *S2.4* Un almacén grande de descuento evalúa cuántos cajeros necesita. Durante un periodo del día cuando la tasa de llegada de los clientes es aproximadamente uno cada 12 minutos, solamente está abierto un cajero. Toma un promedio de ocho minutos atender a un cliente. El tiempo del cajero se ajusta a una distribución normal con desviación estándar de 1.3 minutos. Puede suponerse que el proceso de llegada es un proceso de Poisson. Encuentre las medidas de rendimiento para este sistema. ¿Qué tanto cambiarían sus cálculos si usted asumiera que la distribución de servicio fuera exponencial?

Solución La tasa de llegada es $\lambda = 5$ por hora y la tasa de servicio es $\mu = 60/8 = 7.5$ por hora, dando $\rho = 5/7.5 = 2/3$. Ya que tanto λ como μ se expresan como el número de clientes por hora, debemos expresar la varianza de la distribución del tiempo de servicio en unidades consistentes. La desviación estándar del tiempo de servicio es 1.3 minutos, o $1.3/60 = 0.02167$ hora. Se concluye que la varianza del tiempo de servicio es $0.02167^2 = 4.6944 \times 10^{-4}$ hora^2. (Como alternativa, podríamos haber expresado λ y μ como el número de clientes por minuto.) Se concluye de la fórmula P-K que

$$L_q = \frac{\lambda^2\sigma^2 + \rho^2}{2(1-\rho)} = \frac{(25)(0.0004694) + 4/9}{(2)(1/3)} = 0.6843.$$

Las medidas de desempeño restantes son

$$L = L_q + \rho = 0.6843 + 2/3 = 1.351.$$

$$W_q = L_q/\lambda = 0.13686 \text{ hora (8.21 minutos)}.$$

$$W = W_q + 1/\mu = 0.13686 + 1/7.5 = 0.2702 \text{ hora (16.21 minutos)}.$$

Así, cada cliente debe esperar hacer cola aproximadamente ocho minutos y estar en el sistema aproximadamente dos veces durante ese lapso. En promedio hay menos de un cliente en la cola.

Ahora, suponga que hemos adoptado la hipótesis de que la distribución de servicio es exponencial. Habríamos obtenido las siguientes medidas de desempeño:

$$L = \rho/(1-\rho) = (2/3)/(1/3) = 2.$$

$$L_q = L - \rho = 2 - 2/3 = 4/3.$$

$$W = L/\lambda = 2/5 = 0.4 \text{ hora (24 minutos)}.$$

$$W_q = L_q/\lambda = 1.3333/5 = 0.2667 \text{ hora (16 minutos)}.$$

Vemos que la hipótesis de que la distribución de servicio sea exponencial conduce a errores sustanciales en las medidas de desempeño del sistema. En especial, ¡W_q es 100 por ciento más alto! El modelo M/M/1 es muy sensible a la hipótesis de que el tiempo de servicio es exponencial. Si los resultados de un análisis de colas van a ser exactos, es muy importante verificar cualesquiera hipótesis en relación con la forma de servicio o el proceso de llegadas observando directamente el sistema.

Un interesante caso especial de la cola M/G/1 es cuando la distribución de servicio es determinista (rotulada M/D/1). La fórmula Pollaczec-Khintchine todavía se aplica con el término de la varianza, $\sigma^2 = 0$. Es interesante observar que cuando el número de servidores sobrepasa a uno, no se conocen las expresiones algebraicas para medir el desempeño. De hecho, se dispone de resultados explícitos sólo para unos cuantos casos adicionales. El lector que tenga interés debe consultar una cobertura más completa de las colas, tal como la que ofrecen Gross y Harris (1985) o Kleinrock (1975). Sin embargo, se dispone de los resultados explícitos de *muchas* versiones del modelo de colas M/M. Los ejemplos incluyen sistemas con servicios prioritarios, meterse en la cola y clientes impacientes, sólo para mencionar unos cuantos.

S2.13 LA COLA M/G/∞

Otra versión del problema de colas con distribución general de servicios para el cual hay resultados explícitos es cuando hay un número infinito de servidores. Los clientes llegan al sistema en forma completamente aleatoria de acuerdo a un proceso de Poisson con tasa λ. La distribución del tiempo de servicio es arbitraria con una tasa de servicio μ. En el momento de la llegada, el cliente entra al servicio. Un número infinito de servidores implica que siempre hay un servidor disponible, sin importar cuantos clientes estén en el sistema. Aunque esto pueda parecer poco realista, muchos problemas reales pueden modelarse de esta manera. Ya que no hay una cola de clientes esperando por el servicio, no hay tiempo de espera por el servicio. Así, ambas medidas de desempeño L_q y W_q son cero. Sin embargo, el número de clientes en el

sistema, L, no es cero. Observe que el número de clientes en el sistema es igual al número de servidores ocupados. El interesante resultado es el siguiente:

Resultado: Para la cola M/G/∞ con tasa de llegada λ y tasa de servicio μ, la distribución del número de clientes en el sistema (o el número de servidores ocupados) en estado permanente es de Poisson con tasa λ/μ. Es decir,

$$P\{N = k\} = \frac{e^{-(\lambda/\mu)}(\lambda/\mu)^k}{k!} \qquad \text{para } k = 0, 1, 2, \ldots$$

Éste es un resultado poderoso. Se concluye que tanto la media como la varianza del número de clientes en el sistema en estado estable es λ/μ.

Ejemplo *S2.5*

Una política común del control de inventarios para artículos valiosos es la política de uno por uno, también conocida como $(S - 1, S)$. Esto implica que la acción objetivo es S y que cuando ocurre una demanda, se coloca una nueva orden de una unidad. Suponga que las demandas se generan por un proceso estacionario de Poisson con tasa λ, y que el plazo de entrega para volver a abastecerse es una variable aleatoria con distribución arbitraria con media $1/\mu$. Este problema es exactamente una cola M/G/∞. El número de clientes en el sistema equivale al número de órdenes pendientes, que por el resultado anterior tiene distribución de Poisson con media λ/μ. Si el plazo de entrega está fijo en τ, el número esperado de unidades por ordenar es simplemente $\lambda\tau$. Se concluye que el número esperado de unidades en existencia es $S - \lambda\tau$. Este resultado puede usarse para determinar una expresión de los costos esperados de mantenimiento, existencia y reabastecimiento, que entonces pueden optimizarse mediante la selección de S. (Véase Hadley y Whitin, 1963, p. 212, por ejemplo.)

Ejemplo *S2.6*

La planeación del personal es una función importante en muchas compañías. Considere el departamento de una compañía con un número deseado de 100 plazas. Suponga que los empleados dejan sus plazas según una tasa de 3.4 al mes y que se lleva un promedio de 4 meses para que la compañía ocupe las plazas libres. El análisis de datos pasados muestra que el número de empleados que salen de la compañía por mes tiene distribución de Poisson y que el tiempo requerido para ocupar las plazas sigue la distribución Weibull. ¿Cuál es la probabilidad de que haya más de 15 plazas sin ocupar para cualquier instante de tiempo? ¿Cuántas plazas se ocupan en promedio dentro del departamento? ¿Cuántas plazas debe tener la compañía con objeto de que la planta tenga un promedio de 100 trabajadores?

Solución

Para determinar la distribución de las plazas, modelamos el problema como una cola M/G/∞. Cada vez que un empleado sale, su plaza entra en la cola de plazas sin ocupar. Suponiendo que la búsqueda del reemplazo comienza inmediatamente, el modelo correcto es un número infinito de servidores. De acuerdo con la teoría, el número esperado de plazas sin ocupar es independiente de la distribución de tiempo que se necesita para reemplazar a cada empleado. El número esperado de plazas sin ocupar es λ/μ. Para esta aplicación, λ corresponde a la tasa de despido de los empleados, que es de 3.4 al mes, y μ la tasa de ocupación de las plazas, que es 1/4 al mes. Así, el número medio de plazas sin ocupar es $\lambda/\mu = 3.4/(1/4) = (3.4)(4) = 13.6$. Entonces hay $100 - 13.6 = 86.4$ plazas ocupadas en promedio.

La probabilidad de que haya más de 15 plazas sin ocupar equivale a que una variable aleatoria de Poisson con media 13.6 sobrepase a 15. Con la interpolación en la tabla A-3, esta probabilidad es aproximadamente 0.29.

También se concluye que, si se asignaran 114 plazas al departamento en vez de 100, habría en promedio 100 plazas ocupadas para cualquier instante de tiempo (aunque el número real es una variable aleatoria).[5]

[5] Agradezco a John Peterson de Smith-Kline-Beecham por incitarme a considerar esta aplicación.

S2.14 OPTIMIZACIÓN DE LOS SISTEMAS DE COLAS

El análisis clásico de las colas es más descriptivo que prescriptivo. En la práctica, esto significa que las diferentes distribuciones de entrada y de servicio se determinan según las medidas de desempeño. Estas medidas no se traducen directamente en decisiones óptimas relativas al diseño del sistema. Esta sección muestra lo que habría que hacer con miras a desarrollar modelos para determinar la configuración óptima de un sistema de colas.[6]

Consideremos algunos problemas típicos de diseño que surgen en los sistemas de servicio de colas y lo que se haría con los resultados de la teoría de colas para determinar las configuraciones del sistema óptimo.

Problemas típicos de diseño del sistema de servicio

1. La junta estatal de carreteras necesita determinar el número de casetas que deben estar disponibles en un nuevo camino de cuota interestatal. Entre más casetas estén abiertas para cualquier instante de tiempo, menos tiempo tendrán que esperar los usuarios. Sin embargo, agregar casetas requiere un costo extra de una sola vez para su construcción y costos adicionales progresivos de salarios para quien atienda la caseta.

2. Un importante fabricante de unidades de disco duro de Winchester está construyendo una planta para producir cabezas de lectura y escritura. La gerencia de la compañía está considerando varias opciones para el equipo de fabricación. Una nueva máquina para enrollar las bobinas electrónicas ha duplicado la producción del equipo convencional, pero ha más que triplicado el costo. ¿Se justifica la inversión?

3. Un servicio de procesamiento de palabras está considerando el tamaño de la base de clientes que va a desarrollar. La compañía desea tener un número suficientemente grande de clientes para mantenerse ocupada, pero no tantos como para que no pueda suministrar un tiempo de turnos razonable.

Marco de trabajo de modelado

1. Considere el ejemplo de la junta estatal de carreteras. Entre más tiempo pasen los usuarios en la carretera, menos tiempo pasarán trabajando y contribuyendo a la sociedad. Si vemos esta meta como una optimización de la sociedad, entonces resulta evidente que existe un beneficio económico directo en reducir tiempo de traslado. Suponga que un análisis económico del problema de la carretera condujo a estimar el costo en que se incurre cuando un usuario espera w unidades de tiempo como la función $h(w)$. Un caso típico se ilustra en la figura S2-8.

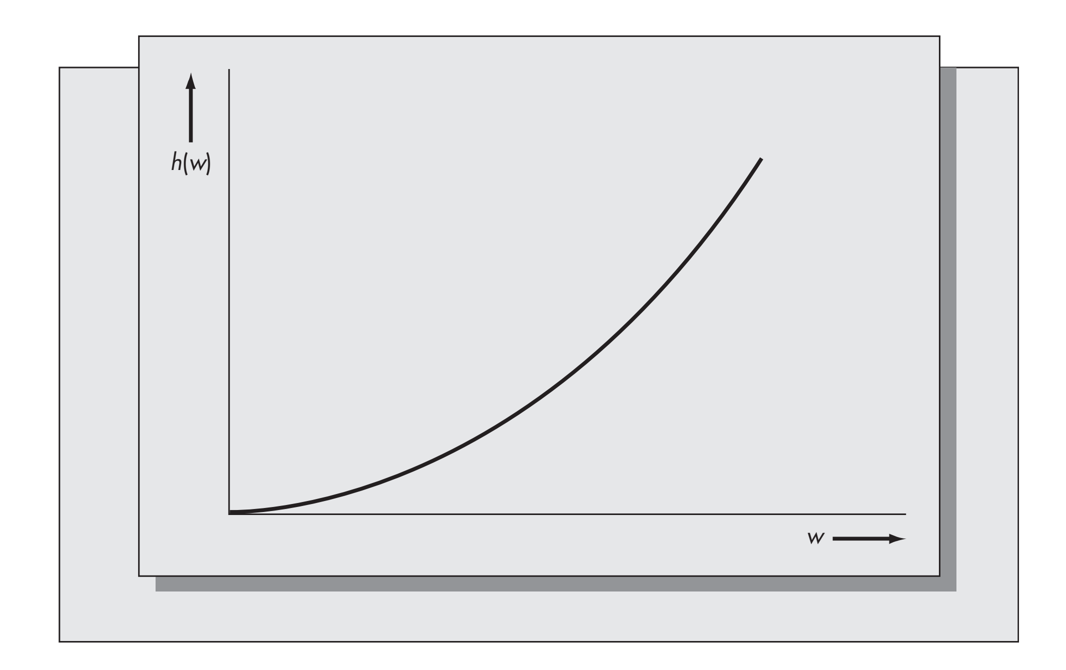
Sea W el tiempo de espera de un cliente escogido aleatoriamente. Entonces W es una variable aleatoria. Para la cola M/M/1, mostramos que W tiene distribución exponencial con parámetro $\mu/(\mu - \lambda)$. Dada la distribución de W, se concluye que la esperanza del costo de espera de un cliente escogido aleatoriamente es

$$E(h(W)) = \int_0^\infty h(w)f(w)\,dw.$$

Ya que la tasa de llegada de los clientes es λ unidades por unidad de tiempo, el costo de espera total por unidad de tiempo es $\lambda E(h(W))$. En el caso de las casetas, se determinaría la distribución de W para cada número de servidores que se consideran, digamos W_c. Si el costo por

[6] Los resultados de esta sección se basan en el capítulo 17 de Hillier y Lieberman (1990).

FIGURA S2-8
Una función típica del costo de espera.



unidad de tiempo para mantener a cada servidor es S, entonces el objetivo sería encontrar el valor óptimo de c para minimizar

$$cS + \lambda E h(W_c).$$

2. Considere el ejemplo de la compañía que produce cabezas de lectura y escritura para las unidades de disco duro. La compañía puede comprar un μ más grande (la tasa de servicio), pero sólo a un costo mayor. Para determinar la mejor decisión en este caso, la compañía debería tener la capacidad de cuantificar los costos asociados con los diferentes niveles de servicio. Suponga que el costo anual de la operación de manufactura cuando la tasa de producción es μ está dado por la función $f(\mu)$. Ya que el costo disminuye a medida que aumenta μ, ésta sería una función monotónicamente decreciente de μ. Además, suponga que el costo de compra de equipo al contado para la tasa de servicio μ es $C(\mu)$. Esperaríamos que $C(\mu)$ fuera una función de μ monotónicamente creciente. Sea I la tasa de interés anual de las inversiones alternas. Entonces el costo anual total es

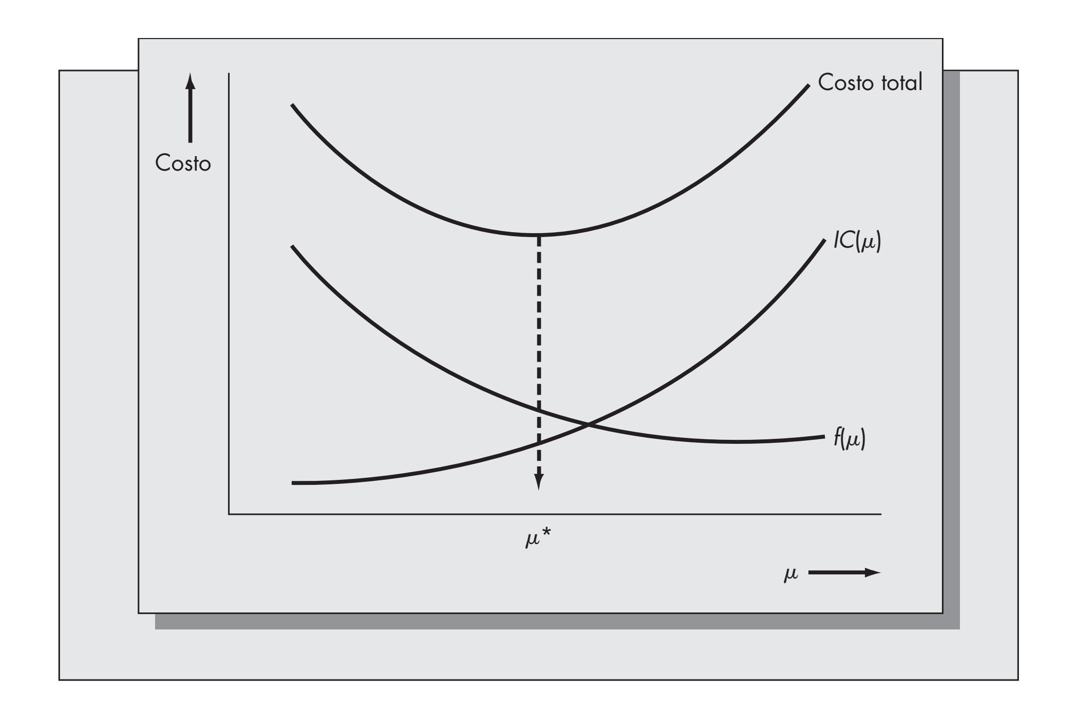
$$IC(\mu) + f(\mu).$$

Esta función será convexa en μ (véase la figura S2-9), de modo que exista un valor de μ óptimo de minimización que pueda encontrarse fácilmente. Cuando existan solamente muchos valores posibles de μ, la función objetivo puede evaluarse para estos valores y es posible tomar la decisión que arroje el costo más bajo.

3. Considere el ejemplo del servicio de procesamiento de palabras. En este caso la variable de decisión es la tasa de llegada λ. Entre más grande sea la base de clientes, la compañía recibirá más trabajos y será mayor el valor de λ. Existen varias formulaciones posibles para este problema. Una sería determinar el valor del número esperado en el sistema que permita a la compañía cumplir sus obligaciones. En ese caso, supondríamos que c y μ están dados y que el objetivo es determinar λ de modo que L será igual a un valor objetivo. Otro enfoque sería encontrar λ para que la probabilidad de que el número de clientes en el sistema no sobrepase a un nivel objetivo sea cuando menos alguna probabilidad específica (tal como .95).

Para ilustrar este caso, considere el siguiente ejemplo específico. Mary Worth dirige su propio servicio de procesamiento. Ella requiere un promedio de 1.2 horas para terminar un

FIGURA S2-9
Optimización de la tasa de servicio μ.



trabajo, pero el tamaño de los trabajos varía considerablemente y es posible describirlo mediante una distribución exponencial. Además, el patrón de llegada de los trabajos resulta ser completamente aleatorio. Mary trabaja ocho horas al día y no desea tener más de dos días de trabajo acumulados en cualquier instante de tiempo. Suponga que ella quiere que la posibilidad de que ocurra esto no sobrepase el 5 por ciento. Esto implica que el número de clientes en la cola no debe sobrepasar a $16/1.2 = 13.33$. Entonces, Mary desea una tasa de llegada de los trabajos de modo que

$$P\{L > 13\} \leq 0.05.$$

Demostramos en la solución del ejemplo S2.1 que

$$P\{L > k\} = \rho^{k+1.}$$

Entonces, buscamos despejar el valor de λ

$$(\lambda/\mu)^{13} = 0.05.$$

Usando $\mu = 1/1.2 = 0.8333$ por hora, y calculando los logaritmos naturales de ambos miembros, obtenemos

$$13 \ln(\lambda/0.8333) = \ln(0.05)$$
$$\ln(\lambda/0.8333) = -0.23044$$
$$\lambda/0.8333 = \exp(-0.23044) = 0.7942,$$

dando $\lambda = 0.662$.

Entonces, la solución es que ella debe planear tener suficientes clientes para generar aproximadamente 0.662 trabajo cada hora, o aproximadamente 5.3 trabajos por día.

Asignar los territorios de servicio a los mecánicos, o decidir qué región debe cubrir un banco de sangre, son otros ejemplos de problemas en los cuales el objetivo es hallar el valor óptimo de la tasa de llegada, λ.

S2.15 SIMULACIÓN DE LOS SISTEMAS EN COLA

Muchos problemas de colas encontrados en el mundo real no son asequibles a los tipos de análisis matemáticos que se estudian en este suplemento. Algunas de las razones son:

1. El problema incluye una red de colas. Las redes de colas son comunes en los sistemas de manufactura donde la salida de un proceso es la entrada de otro o de otros procesos. Algunos sistemas, con bucles de retroalimentación y otras características poco comunes, son muy complejos. Generalmente estos sistemas son demasiado complicados para analizarse matemáticamente.
2. El interés radica en el comportamiento transitorio en vez de en el de estado estable. Los resultados matemáticos estudiados aquí sólo son para el segundo caso. Estado estable significa que el sistema evoluciona por un lapso suficientemente largo de modo que las condiciones iniciales no afectan la distribución de probabilidad de los estados del sistema. En muchos problemas reales, el comportamiento a corto plazo es una parte importante de una administración apropiada de los sistemas.
3. Las distribuciones de probabilidad no exponenciales de las llegadas o de los servicios y/o de características o restricciones poco comunes pueden imposibilitar el análisis. Se conoce poco sobre los sistemas de colas G/G/*c*, especialmente con una sala de espera finita, un servicio prioritario, etc. Aun para colas con un solo servidor en estado estable, los problemas de colas son demasiado complejos para analizarse matemáticamente.

¿Cuáles son las alternativas disponibles si un problema de colas no es uno de aquellos para los cuales se conocen los resultados? La respuesta es la simulación. Una simulación (en nuestro contexto) es un programa de computadora que recrea los pasos esenciales del escenario de un problema. En cierto sentido, la computadora "experimenta" el problema. De esta manera, puede verse cómo el sistema responde a diferentes valores paramétricos disponibles para el diseñador del sistema. Las simulaciones evolucionan con el tiempo. Los sistemas pueden simularse durante años en el espacio de unos cuantos segundos en una computadora.

Los problemas asequibles a la simulación la mayoría de las veces incluyen algún elemento de aleatoriedad. Los simuladores basados en computadora que incluyen a la aleatoriedad se llaman simuladores de Monte Carlo. En el corazón de una simulación de Monte Carlo radican los números aleatorios.

Los números aleatorios en este contexto se extraen de una distribución uniforme (0, 1). Es decir, son números entre cero y uno con la propiedad de que cada número extraído tiene la misma posibilidad de ser seleccionado. Los generadores de números aleatorios son algoritmos que producen lo que resulta ser ocurrencias independientes de variables uniformes. Los algoritmos usados no recorren ciclos durante un gran número de pasos, produciendo así secuencias numéricas que resultan ser aleatorias. Sin embargo, ya que los algoritmos recursivos que se emplean son deterministas, la cadena resultante de números se conoce como número "pseudoaleatorio". (Véase Fishman, 1973, por ejemplo.)

A partir de variables uniformes se obtienen observaciones de variables aleatorias que virtualmente poseen cualquier distribución que use resultados de la teoría de la probabilidad. Por ejemplo, el teorema del límite central dice que las sumas de las variables aleatorias independientes tiene una distribución aproximadamente normal. (Éste es un enunciado *muy* general del teorema del límite central.) Por lo tanto, la suma de un número razonable de extracciones independientes de una distribución uniforme (0, 1) será aproximadamente normal. (La convergencia se presenta con mucha rapidez, de modo que el número no tiene que ser muy grande.) Por ejemplo, si hacemos que U_1, U_2 . . . sean extracciones sucesivas de una distribución uniforme (0, 1) entonces

$$Z = \sum_{i=1}^{12} U_i - 6$$

es aproximadamente normal estándar. (Esto tiene solamente un objetivo ilustrativo. En la práctica hay formas más eficientes de generar variables normales.)

La simulación puede usarse para analizar muchos tipos de problemas complejos con incertidumbre, pero probablemente se ha usado con más frecuencia para los problemas de colas. Los simuladores pueden escribirse en código fuente (tal como Fortran o C), o construirse con el uso de paquetes de simulación de uso especial, tales como GPSS, Simscript, o SLAM. Muchos paquetes están diseñados específicamente para ciertos tipos de aplicaciones tales como problemas de colas o de redes. Swain (1995) suministra una revisión detallada de más de 50 paquetes diseñados para la simulación de eventos discretos en computadoras personales.

Más recientemente, se han hecho populares dos enfoques del desarrollo de simulaciones, abriéndose así el uso de la simulación a una audiencia mucho más numerosa. Los programas de hoja de cálculo han ganado muchísima popularidad en años recientes. Ya que muchas hojas de cálculo incorporan un generador de números aleatorios, pueden usarse para construir simuladores. Excel, específicamente, puede generar observaciones a partir de distribuciones numerosas. Para facilitar la construcción de simuladores, puede considerarse el @RISK aditivo, que incluye un arreglo mucho más amplio de distribuciones y una generación conveniente de los reportes. (Los ejemplos con el uso de @RISK pueden encontrarse en Winston, 1996).

Para redes de colas complejas, los paquetes de simulación basados en gráficos son mucho más fáciles de usar que las hojas de cálculo. Estos programas permiten al usuario construir un modelo del sistema usando íconos gráficos. Los íconos representan las instalaciones de servicio y las áreas de espera y las flechas representan la dirección del flujo. Las llegadas aleatorias y los servicios aleatorios pueden incorporarse fácilmente en el modelo. Un producto basado en Windows es ProModel, desarrollado y vendido a través de la ProModel Corporation. Con un paquete como éste, los usuarios experimentados pueden construir simulaciones de sistemas complejos muy rápidamente. El programa emplea la animación en vivo para mostrar la simulación en acción, y se recopilan estadísticas resumidas después que la simulación ha seguido su curso.

Un comentario final. Aunque hemos estudiado la simulación solamente en el contexto de las colas, es una herramienta poderosa para resolver muchos tipos de problemas complejos. Las simulaciones se han usado para problemas de manejo de inventarios, programación, planificación del tráfico, asignación de recursos, modelación financiera, etc. Sin embargo, a juzgar por los productos de software disponibles, las aplicaciones de colas resultarían ser las más comunes.

Bibliografía

Feller, W. *An Introduction to Probability Theory and Its Applications*. Vol. 2. Nueva York: John Wiley & Sons, 1966.

Fishman, G. *Concepts and Methods in Discrete Event Digital Simulation*, Nueva York: John Wiley & Sons, 1973.

Gross, D. y C. M. Harris. *Fundamentals of Queuing Theory.* 2a. ed. Nueva York: John Wiley & Sons, 1985.

Hadley, G. y T. M. Whitin. *Analysis of Inventory Systems.* Englewood Cliffs, NJ: Prentice Hall, 1963.

Hillier, F. S. y G. J. Lieberman. *Introduction to Operations Research.* 5a. ed. Nueva York: McGraw-Hill, 1990.

Kleinrock, L. *Queuing Systems.* Vol. 1, *Theory.* Nueva York: Wiley Interscience, 1975.

Kleinrock, L. *Queuing Systems.* Vol. 2, *Computer Applications.* Nueva York: Wiley Interscience, 1976.

Swain, J. J. "Simulation Survey: Tools for Process Understanding and Improvement". *OR/MS Today* 22, núm. 4, agosto de 1995, pp. 64-79.

Winston, W. L. *Simulation Modeling Using @ RISK.* Belmont, CA: Duxbury Press, 1996.

Capítulo **Nueve**

Programación de proyectos

Panorama general del capítulo

Objetivo

Entender cómo se usan las técnicas matemáticas y gráficas para ayudar a programar proyectos complejos en una organización.

Puntos clave

1. *Representación del proyecto e identificación de la ruta crítica.* Se cuenta con dos técnicas gráficas convenientes para representar un proyecto. Una es la gráfica de Gantt. Ésta se usó en el capítulo 8 para representar programas de secuencia en máquinas múltiples. Sin embargo, representar un proyecto con un diagrama de Gantt tiene una desventaja importante. No se muestran las relaciones de precedencia (es decir, la especificación de cuáles actividades deben preceder a otras). Para superar este inconveniente, representamos el proyecto con una red en lugar de un diagrama de Gantt. Una red es un conjunto de nodos y arcos dirigidos. Los nodos corresponden a hitos en el proyecto (la terminación de algún subconjunto de actividades) y los arcos a las actividades específicas.

   El objetivo consiste en identificar la ruta crítica o más larga en la representación de red. Siguiendo la idea de que "una cadena es sólo tan fuerte como su eslabón más débil", un proyecto no puede acabar hasta que se terminen todas las actividades a lo largo de la ruta crítica. La longitud de ésta da el tiempo más corto de terminación del proyecto. Las actividades que no estén a lo largo de la ruta crítica (actividades no críticas) tienen tiempos de holgura: es decir, pueden retrasarse sin necesariamente retrasar al proyecto. En la sección 9.2, presentamos un algoritmo para identificar la ruta crítica en una red. (Éste constituye sólo uno de varios métodos de solución.)
2. *Métodos de costeo de tiempo.* Piense en un proyecto de construcción. Cada día adicional que transcurra elevará los costos. Éstos incluyen los costos directos de la mano de obra del personal que participa en el proyecto, los costos asociados con el uso de equipo y materiales, y los costos generales. Supongamos que se tiene la opción de disminuir el tiempo de las actividades seleccionadas, pero también con algún costo. A medida que disminuyen los tiempos requeridos para las actividades a lo largo de la ruta crítica, aumentan los costos de agilización pero disminuyen los proporcionales a la duración del proyecto. Entonces, existe una duración óptima del proyecto que balancea estos dos costos competidores. El problema de la optimización del costo para la duración de un proyecto se puede resolver de manera manual o por medio de la programación lineal.

3. *Programación de proyectos con tiempos inciertos de actividades.* El tiempo requerido para realizar tareas específicas, como en el caso de algunos proyectos de construcción, puede predecirse con exactitud. En la mayoría de los casos, es posible usar la experiencia pasada como una guía exacta, incluso para nuevos proyectos. Sin embargo, éste no es el caso con los proyectos de investigación. Cuando se busca la solución de un problema no resuelto, o se diseña una pieza de equipo completamente nueva, resulta difícil, si no imposible, predecir con exactitud los tiempos de las actividades. Una hipótesis más razonable es que los tiempos de las actividades son variables aleatorias con alguna distribución específica.

   La técnica de evaluación y revisión de proyectos (PERT, *project evaluation and review technique*) es un método que admite abiertamente los tiempos de las actividades inciertas. Esta técnica fue desarrollada por la Marina para ayudar a planear el proyecto del submarino Polaris en 1958. El enfoque PERT consiste en suponer que los planificadores especifican un tiempo mínimo, a, un tiempo máximo, b, y un tiempo más probable, m, para cada actividad. Estas estimaciones se usan con el fin de construir una distribución beta para cada tiempo de actividad. La hipótesis PERT es que la ruta crítica será aquélla con el mayor tiempo esperado de terminación (que no necesariamente es el caso), y el tiempo total del proyecto será la suma de los tiempos a lo largo de la ruta crítica. Suponiendo que los tiempos de las actividades son variables aleatorias independientes, se calculan la media y la varianza a lo largo de la ruta crítica sumando las medias y las varianzas de los tiempos de las actividades. Entonces se usa el teorema del límite central para justificar la hipótesis de que el tiempo de terminación del proyecto tiene distribución normal con media y varianza calculadas, como ya se describió. Observe que ésta es sólo una aproximación, puesto que no hay ninguna garantía de que la ruta con el mayor tiempo esperado de terminación resulte ser la ruta crítica. Determinar la verdadera distribución del tiempo de terminación del proyecto resulta ser un problema en general muy difícil. Sin embargo, PERT suministra una aproximación razonable y ciertamente es una mejora sobre el método de la ruta crítica (CPM, *critical path method*).

4. *Consideraciones de los recursos.* Piense en un departamento dentro de una compañía en la que se están desarrollando de manera simultánea varios proyectos. Suponga que cada miembro del departamento trabaja en más de un proyecto a la vez. Como el tiempo de cada trabajador es limitado, cada gerente de proyectos está compitiendo por un recurso limitado, es decir, el tiempo de los trabajadores. Podrían imaginarse otros casos en los cuales el recurso limitado puede ser una pieza de equipo, tal como una sola supercomputadora en una compañía. En estos casos, incorporar las restricciones de los recursos en la función de planeación del proyecto puede ser un verdadero reto. A continuación se presenta un ejemplo de balance de recursos, que no constituye un método de propósito general para dar solución a este problema.

Roma no se construyó en un día, ni tampoco las pirámides de Egipto, el edificio Empire State, el puente Golden Gate o la torre Eiffel. Todos estos fueron proyectos complejos que necesitaron una planeación y una coordinación cuidadosas. ¿Cómo se organizan y monitorean proyectos masivos de este tipo? Una administración efectiva podría cristalizar o malograr un proyecto. Si bien se requieren muchos talentos distintos para ser un gerente de proyectos eficiente, las técnicas cuantitativas pueden ser de gran ayuda. Dichas técnicas son el tema de este capítulo.

¿Cuáles son las consecuencias de una mala administración de proyectos? Una es el exceso de gastos. ¿Cuántas veces hemos oído que los miembros del Congreso expresan su amargura por el exceso de gastos en los proyectos militares? En algunos casos este exceso no puede evitarse, ya que surgen obstáculos imprevistos o no se pueden resolver problemas tecnológicos tan fácilmente como se espera. Estos problemas, sin embargo, son consecuencia de una deficiente programación y administración de proyectos en muchas ocasiones.

Los más vulnerables a las demoras y al exceso de gastos son quizás los proyectos grandes y complejos que involucran a los gobiernos en asociación con las empresas. Uno de estos casos es el Trans-Alaska Pipeline System (Sistema de tuberías trans-Alaska), diseñado para transportar grandes cantidades de petróleo desde la bahía Prudhoe en la vertiente norte de Alaska, hasta el Puerto Valdez en el golfo del mismo estado. Goodman (1988) lo describe como uno de los proyectos de diseño y construcción más complejos y masivos de los últimos tiempos. Las barreras políticas, las preocupaciones ambientales y las disputas por los contratos atormentaron al proyecto desde el inicio. La lista de jugadores importantes cambió varias veces, lo que condujo a que algunas compañías no tuvieran suficiente tiempo (especialmente Bechtel) para hacer un buen trabajo de planeación en sus proyectos. La Alyesca Pipeline Service Company, responsable de gran parte de la construcción actual de la tubería, incurrió en demasiados costos excesivos, en parte debido a una mala administración de proyectos. En retrospectiva, es evidente que muchos de los problemas surgieron como consecuencia de que no hubo un solo equipo que supervisara todo el ciclo del proyecto integrado.

Los métodos de programación y administración de proyectos han constituido una parte importante de la negociación para muchas compañías. Por ejemplo, la Lockheed-Martin Missiles and Space Company de Mountain View, California, usa métodos de programación no sólo para llevar a cabo el monitoreo y el control de los proyectos, sino también para preparar licitaciones y desarrollar propuestas. Lockheed, de hecho, formó parte del equipo que desarrolló PERT, una técnica que se estudia de manera exhaustiva en este capítulo.

Las técnicas analíticas de la administración de proyectos se revisan en este capítulo. Una administración efectiva del personal puede ser un factor igualmente importante para lograr que los proyectos se hagan a tiempo y dentro del presupuesto. La compañía debe crear una estructura y un ambiente que conduzcan a una motivación apropiada de los empleados. Las imperfecciones en el diseño organizacional y en las estructuras de incentivos pueden ser un problema tan grave como una planificación deficiente del proyecto.

Los métodos de administración considerados en este capítulo se han usado para planear proyectos a largo plazo, como el lanzamiento de nuevos productos, la organización de proyectos de investigación y la construcción de instalaciones de producción. Los métodos también se han empleado en proyectos más pequeños como la edificación de casas residenciales. Dos técnicas que se tratan con detalle son el *método de la ruta crítica* (CPM) y la *técnica de evaluación y revisión de proyectos* (PERT). Ambos métodos se desarrollaron casi al mismo tiempo, a finales de la década de los cincuenta, para resolver problemas de administración de proyectos muy diferentes. Aunque actualmente las dos etiquetas se usan en forma intercambiable, vamos a conservar la terminología consistente con el objetivo original de los métodos. Es decir, CPM se ocupa de problemas puramente determinísticos, mientras que PERT permite la aleatoriedad de los tiempos en las actividades.

Los elementos primarios del análisis de ruta crítica son:

1. *La definición del proyecto*. Un planteamiento preciso del proyecto, las metas que persigue y los recursos y el personal que requiere.

2. *Definición de las actividades*. El proyecto se debe separar en un conjunto de tareas o actividades indivisibles. El diseñador del proyecto debe especificar el contenido de trabajo y estimar los tiempos de actividad. En muchas ocasiones, la parte más difícil de la planeación de proyectos reside en encontrar la mejor manera de separar el proyecto en una serie de actividades diferentes.

3. *Relación entre actividades*. Una parte importante de la metodología de la planeación de proyectos consiste en especificar las interrelaciones entre las actividades. Éstas describen la secuencia lógica para terminar las actividades que conforman el proyecto y se les conoce como *restricciones de precedencia*.

4. *Programación del proyecto*. Un calendario de proyecto es una relación clara de los tiempos de inicio y de terminación de todas las actividades que lo constituyen. En este capítulo se mostrarán técnicas para obtener un programa eficiente al especificar los tiempos de las actividades y las restricciones de precedencia.

5. *Monitoreo del proyecto*. Cuando las actividades ya han sido adecuadamente definidas y se ha determinado un programa, deben ponerse en su lugar los controles apropiados para asegurar que los hitos del proyecto se cumplan. Si surgen problemas imprevistos el director del proyecto debe estar preparado para revisar los programas existentes.

Debido al nivel de precisión y detalle que se requiere, el análisis de ruta crítica es muy efectivo cuando el proyecto puede expresarse fácilmente como un grupo de actividades bien definidas. Los proyectos de construcción se sitúan en esta categoría. Las restricciones de precedencia son directas y no es difícil calcular los tiempos de las actividades. Por esta causa, el CPM ha encontrado una amplia aceptación en la industria de la construcción. Por otra parte, en la literatura también se reportan casos de estudio de aplicaciones exitosas en diferentes ambientes, incluyendo la milicia, el gobierno y actividades no lucrativas.

9.1 REPRESENTACIÓN DE UN PROYECTO COMO UNA RED

En la programación y control del piso de producción, tal como lo vimos en el capítulo 8, un *diagrama de Gantt* es una gráfica de barras horizontales en la que cada actividad corresponde a una barra separada. Considere a continuación este ejemplo sencillo.

Ejemplo *9.1*

Piense en un proyecto conformado por cinco actividades: A, B, C, D y E. La figura 9-1 muestra un programa para llevar a cabo este proyecto en la forma de un diagrama de Gantt. De acuerdo con la figura, la tarea A inicia a las 12:00 y termina a la 1:30. Las tareas B y C inician a la 1:30; B termina a las 2:30, C termina a las 3:30, etc. Finalmente, el proyecto termina a las 6:00.

Observe que la figura 9-1 no da información sobre las relaciones entre las actividades. Por ejemplo, la figura implica que E no puede iniciar hasta que se termine D. Sin embargo, suponga que E inicie en cualquier momento después de la terminación de A, B y C. Entonces, E puede comenzar a las 3:30 y el proyecto se puede terminar a las 5:00 en lugar de a las 6:00.

A pesar de que el diagrama de Gantt es un medio eficaz para representar el programa, no constituye una herramienta de planeación muy útil porque no muestra las restricciones de precedencia. Por esta razón, el proyecto se representa gráficamente como una red. A diferencia de los diagramas de Gantt, las redes muestran explícitamente las restricciones de precedencia.

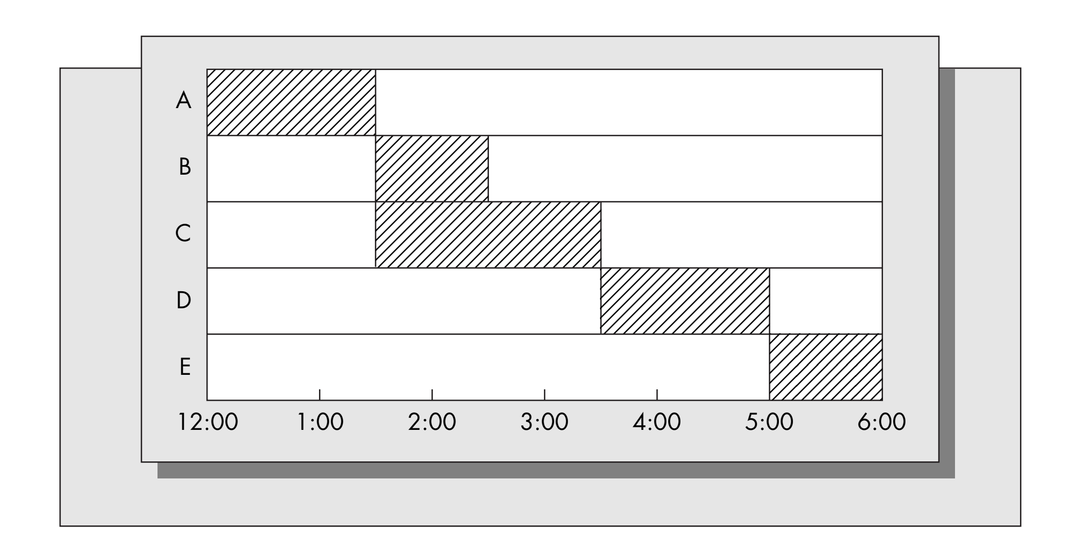
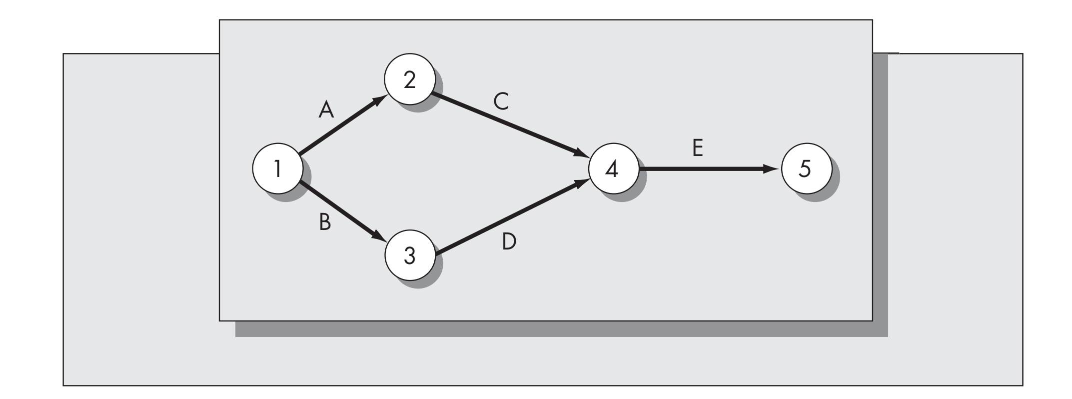


FIGURA 9-1
Diagrama de Gantt para cinco actividades (véase el ejemplo 9.1).

FIGURA 9-2
Red de proyecto para el ejemplo 9.2.



Una red es una colección de nodos y de arcos dirigidos. De acuerdo con la representación tradicional de red de un proyecto, un nodo corresponde a un evento y un arco a una actividad o a una tarea. Los eventos pueden ser (1) el inicio de un proyecto, (2) la terminación de un proyecto o (3) la terminación de algún grupo de actividades. Este método de representación se conoce como *actividades en las flechas* e, históricamente, es el medio más común para representar un proyecto como una red. Un método de representación alternativo es el de actividades en los nodos. Si bien este último método tiene algunas ventajas, rara vez es utilizado. En la sección 9.2 se ilustrará el método de actividades en los nodos.

Ejemplo *9.2*

Piense en un proyecto que supone cinco actividades A, B, C, D y E, las cuales satisfacen las siguientes relaciones de precedencia:

1. Ni A ni B tienen ningunos predecesores inmediatos.
2. A es un predecesor inmediato de C.
3. B es un predecesor inmediato de D.
4. C y D son predecesores inmediatos de E.

La red de este proyecto se muestra en la figura 9-2. El nodo 1 siempre es el inicio del proyecto. Todas las actividades que no posean predecesores inmediatos emanan del nodo 1. Como A es un predecesor inmediato de C, debemos tener un nodo que represente la terminación de A. En forma similar, como B es un predecesor inmediato de D, debe existir un nodo que represente la terminación de B. Observe que, si bien, tanto A como B también deben terminarse antes que E, no son predecesores *inmediatos* de E.

La representación de un proyecto mediante una red no siempre es tan directa.

Ejemplo *9.3*

Considere el ejemplo 9.2, excepto que ahora debe reemplazar a (3) por (3?):

3′. A y B son predecesores inmediatos de D.

Intente encontrar una representación de red para este proyecto. Podría pensarse que la representación correcta es la de la figura 9-3, puesto que esta red implica que D debe esperar por A y B. No obstante, esta representación es incorrecta, pues muestra que C debe esperar también por A y B.

El conjunto de relaciones precedentes (1), (2), (3’) y (4) requieren que se introduzca una pseudoactividad entre los nodos 2 y 3. La representación correcta del sistema se da en la figura 9-4. La pseudoactividad, identificada como P, es un arco dirigido del nodo 2 al nodo 3 y se representa con una línea discontinua. Observe que la dirección de la flecha es importante. Con la pseudoactividad, el nodo 3 corresponde a la terminación de ambas actividades A y B, mientras que el nodo 2 todavía corresponde a la terminación sólo de la actividad A.

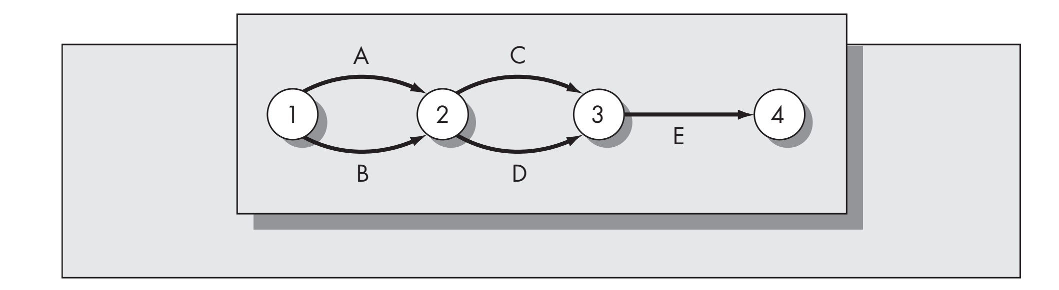


FIGURA 9-3
Representación incorrecta de la red del ejemplo 9.3.

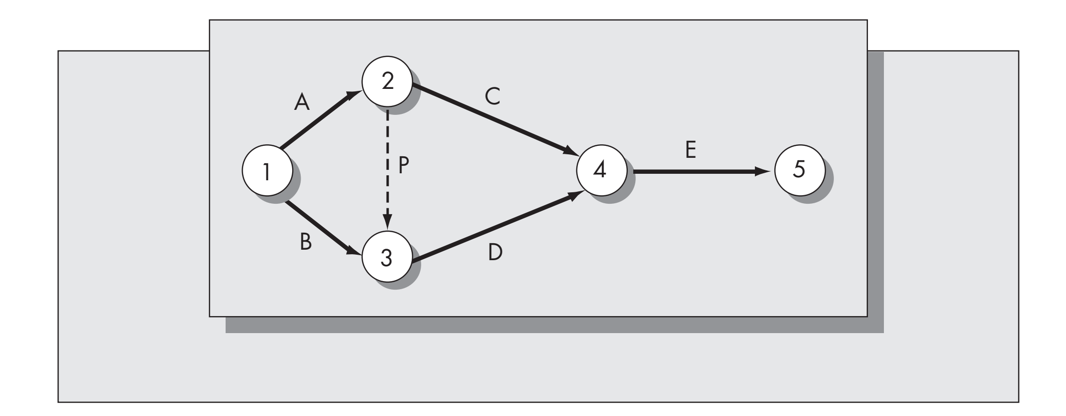


FIGURA 9-4
Representación correcta de la red del ejemplo 9.3.

La definición de un conjunto apropiado de tareas es un aspecto importante del análisis de ruta crítica. Un problema que puede surgir es que se definan las tareas con demasiada imprecisión. Para ilustrar esto, suponga que en el ejemplo 9.3, E podría iniciar solamente cuando terminara la mitad de C. Esto implica que C tendría que separarse aún más en dos actividades adicionales, una que represente la primera mitad de C y otra que represente la segunda mitad de C. Al contrario, es posible que las actividades se definan con demasiada minuciosidad. Esto quizás conduzca a una representación de red demasiado complicada, con partes de la red que se parezcan a la figura 9-3.

9.2 ANÁLISIS DE LA RUTA CRÍTICA

Podemos comenzar a explorar los métodos disponibles para responder a varias preguntas una vez que se ha representado el proyecto como una red. Por ejemplo,

1. ¿Cuál es el tiempo mínimo que se requiere para terminar el proyecto?
2. ¿Cuáles son los tiempos de inicio y de terminación para cada una de las actividades?
3. ¿Qué actividades pueden retrasarse sin aplazar la terminación del proyecto?

Una ruta a través de la red es una secuencia de actividades desde el nodo 1 hasta el nodo final. Una ruta se identifica siguiendo la secuencia de nodos visitados o por la secuencia de arcos (actividades) que se atraviesan. En la red ilustrada por la figura 9-4 hay exactamente tres rutas diferentes: 1-2-4-5 (A-C-E), 1-2-3-4-5 (A-P-D-E) y 1-3-4-5 (B-D-E). Cada ruta es una secuencia de actividades que satisfacen las restricciones de precedencia. Como el proyecto se termina cuando se acaban *todas* las actividades, se concluye que todas las rutas desde el nodo inicial hasta el nodo final deben atravesarse. Así, el tiempo mínimo para terminar el proyecto debe ser el mismo que la longitud de la ruta *más larga* de la red.

Considere el ejemplo 9.3 y suponga que los tiempos de las actividades son los mismos que los indicados en la figura 9-1. Los tiempos de las actividades y las restricciones de precedencia son

Actividad	Tiempo (horas)	Predecesores inmediatos
A	1.5	
B	1.0	
C	2.0	A
D	1.5	A y B
E	1.0	C y D

Las longitudes de las tres rutas son

Ruta	Tiempo requerido para terminar (horas)
A–C–E	4.5
A–P–D–E	4.0
B–D–E	3.5

Las tres rutas deben completarse para terminar el proyecto. La más larga de este ejemplo es obviamente A-C-E. Se le conoce como la *ruta crítica*. Su longitud es el tiempo de terminación *mínimo* del proyecto. Las actividades situadas a lo largo de ésta se conocen como críticas. El retraso de una actividad crítica conduce a una demora del proyecto. Sin embargo, las actividades que no están situadas a lo largo de la ruta crítica tienen cierta holgura. El proyecto no necesariamente se aplaza por la dilación de una actividad no crítica.

En general, la enumeración no es una manera eficiente de encontrar la ruta crítica. Más tarde consideraremos un procedimiento general, que no requiere enumeración, para identificar la ruta crítica y los tiempos de inicio y de terminación para todas las actividades que constituyen el proyecto. Primero describiremos el siguiente caso de estudio.

Ejemplo *9.4*

Simon North e Irving Bonner, consultores de computación, están considerando aventurarse en un proyecto conjunto que incluya el desarrollo de un paquete de software comercial relativamente pequeño para computadoras personales. El programa incluye cálculos científicos para una parte especializada del mercado de ingeniería. North y Bonner han dividido el proyecto en nueve tareas.

La primera tarea consiste en realizar una encuesta de mercado para determinar exactamente qué va a requerir la clientela potencial y cuáles características del software serán probablemente las más atractivas. Una vez que se termine esta etapa, puede comenzar el desarrollo real de los programas. Los requerimientos de programación se sitúan en dos amplias categorías: los gráficos y el código fuente. Como el sistema va a ser interactivo y manejado con iconos, de entrada hay que identificarlos y diseñarlos. Después de que los programadores terminen el diseño de los iconos, pueden proseguir con la segunda parte del desarrollo de los gráficos: el diseño de las pantallas de entrada/salida. Éstas incluyen los diferentes menús y los generadores de reportes requeridos en el sistema.

La segunda parte del proyecto supone codificar los módulos que realizan los cálculos científicos. El primer paso consiste en desarrollar un diagrama de flujo detallado del sistema. Después, los programadores pueden comenzar a trabajar en los módulos. Hay un total de cuatro módulos. El trabajo de los módulos 1 y 2 puede comenzar inmediatamente después de terminar el diagrama de flujo. El módulo 3 requiere partes del módulo 1, de modo que el trabajo en aquél no puede comenzar hasta que se complete el módulo 1. No es posible iniciar la programación del módulo 4 hasta que se terminen los módulos 1 y 2. Una vez terminada la parte gráfica del proyecto así como los módulos, las dos partes

separadas del sistema requieren fusionarse, y el sistema completo necesita probarse y someterse a pruebas de escritorio.

North ha logrado obtener un financiamiento para el proyecto, pero su fuente requiere que el programa esté terminado y listo para el mercado en 25 semanas. Para determinar si esto es factible, los dos programadores dividieron el proyecto en nueve tareas indivisibles y han estimado el tiempo requerido para cada tarea. La lista de tareas, los tiempos que requieren, y sus relaciones de precedencia, se presentan a continuación.

Tara	Tiempo requerido (semanas)	Predecesores inmediatos
A. Realizar la encuesta de mercado	3	
B. Diseño de los iconos gráficos	4	A
C. Desarrollo del diagrama de flujo	2	A
D. Diseño de las pantallas de entrada/salida	6	B, C
E. Codificación del módulo 1	5	C
F. Codificación del módulo 2	3	C
G. Codificación del módulo 3	7	E
H. Codificación del módulo 4	5	E, F
I. Fusión de los módulos y de los gráficos y pruebas del programa	8	D, G, H

El total de tiempo de tareas es de 43 semanas. Con base en esto, los programadores concluyen que sería imposible finalizar su proyecto en las 25 semanas requeridas. Afortunadamente, ven el error de planeación antes de romper relaciones con su fuente de financiamiento. Como algunas de las actividades pueden hacerse en forma concurrente, el proyecto deberá tomar menos de 43 semanas. La representación de red de este proyecto aparece en la figura 9-5. Con el objetivo de satisfacer las restricciones de precedencia requerimos dos pseudoactividades, P_1 y P_2. Determinaremos la ruta crítica sin enumerar todas las rutas a través de la red. Los cálculos de la ruta crítica dan como resultado tanto la ruta crítica como la holgura permisible para cada actividad.

Con el propósito de ilustrarlo, representamos la red de proyecto que incluye el uso del formato con actividades en los nodos de la figura 9-6. Una ventaja de este formato es que no se requieren pseudoactividades, aunque ciertas redes pueden exigir nodos ficticios de inicio y de terminación. Dada una lista de actividades y de los predecesores inmediatos, debe ser fácil encontrar la representación de la red en cualquiera de los dos formatos.

FIGURA 9-5
Red para el ejemplo 9.4.

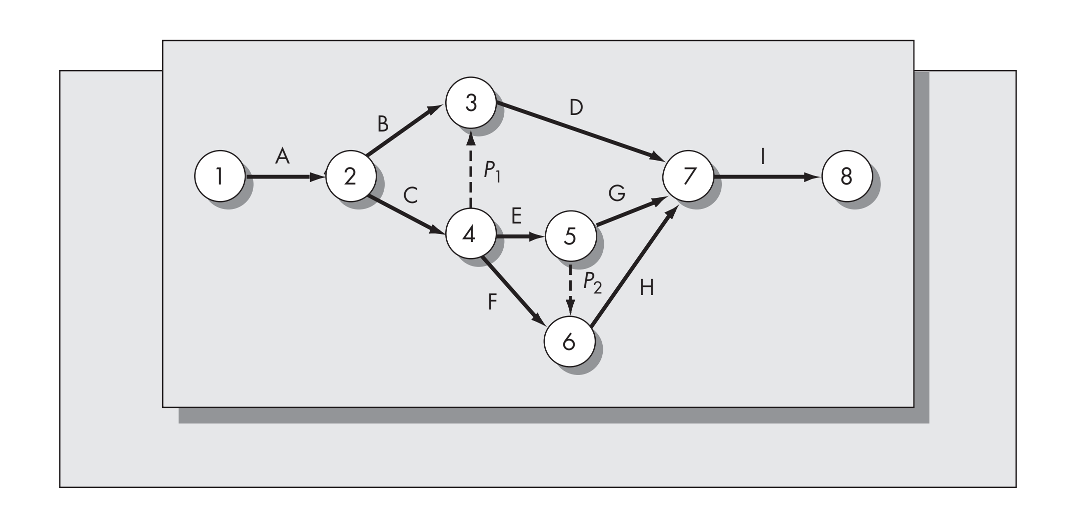
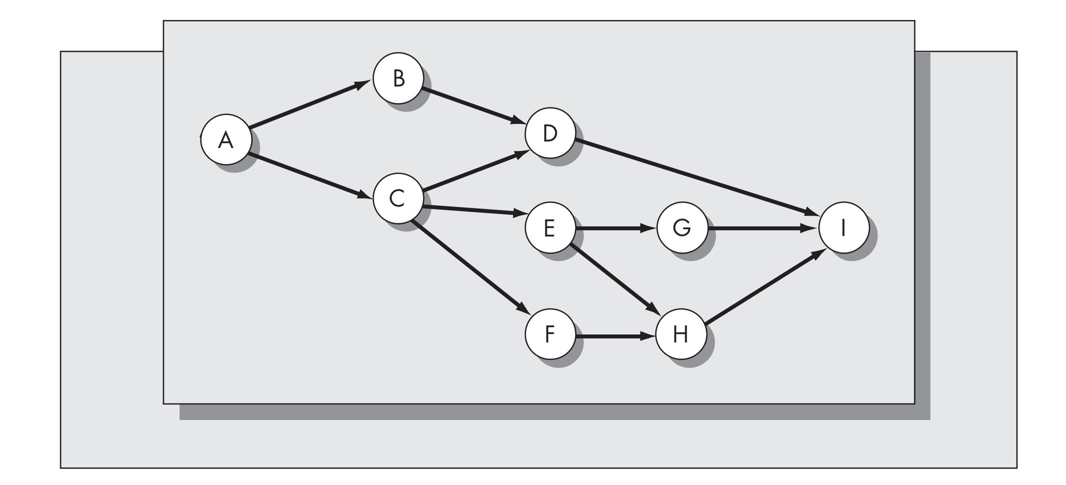


FIGURA 9-6
Red para el ejemplo 9.4: representación con actividades en los nodos.



Por considerar que es una forma más intuitiva de representar el proyecto los profesionales prefieren el método con actividades en las flechas. Lo usaremos exclusivamente para el resto del capítulo.

Determinación de la ruta crítica

Calculamos cuatro cantidades para cada actividad:

ES_i = Tiempo de inicio más próximo para la actividad *i*,

EF_i = Tiempo de terminación más próximo para la actividad *i*,

LF_i = Tiempo de terminación más lejano para la actividad *i* (sin retrasar el proyecto),

LS_i = Tiempo de inicio más lejano para la actividad *i* (sin retrasar el proyecto).

Suponga que t_i representa el tiempo requerido para terminar la actividad *i*. Entonces resulta fácil ver que

$$EF_i = ES_i + t_i$$

y

$$LS_i = LF_i - t_i.$$

Los pasos del proceso son

1. *Calcular los tiempos más próximos para cada actividad*. Se calculan los tiempos más próximos mediante un recorrido hacia adelante a través de la red; es decir, los cálculos se hacen desde el nodo 1 hasta el nodo final.
2. *Calcular los tiempos más lejanos de cada actividad*. Se calculan los tiempos más lejanos mediante un recorrido hacia atrás a través de la red; es decir, los cálculos se hacen desde el nodo final hasta el nodo 1.

Usaremos el ejemplo 9.4 para ilustrar los cálculos. El primer paso es igualar a cero los tiempos de inicio más próximos de todas las actividades que emanan del nodo 1 y sumar los tiempos de las actividades para obtener los tiempos de terminación más próximos. En este caso, la única actividad que emana del nodo 1 es la actividad A. Como B tiene a A como un predecesor inmediato, se concluye que $ES_B = EF_A$. De manera similar para C, $ES_C = EF_A$. Calcule los tiempos de terminación más próximos para B y C sumando los tiempos de las actividades a

los de inicio más próximos: $EF_B = ES_B + t_B$, y $EF_C = ES_C + t_C$. Al resumir los cálculos hasta este punto, tenemos

Actividad	Tiempo	Predecesor inmediato	ES	EF
A	3	—	0	3
B	4	A	3	7
C	2	A	3	5

Resulta más difícil calcular el tiempo de inicio más próximo para D. La actividad D tiene dos predecesores inmediatos, B y C. Eso implica que D no puede iniciar hasta que *tanto* B *como* C hayan sido terminadas. Se concluye que el tiempo de inicio más próximo para D es el *último* de los tiempos de terminación más próximos para B y C. Es decir,

$$ES_D = \text{máx}(EF_B, EF_C) = \text{máx}(7, 5) = 7.$$

Regla general: El tiempo de inicio más próximo de una actividad es el *máximo* de los tiempos de terminación más próximos de sus predecesores inmediatos.

Continuando de esta manera, obtenemos los siguientes tiempos más próximos para todas las actividades restantes:

Actividad	Tiempo	Predecesor inmediato	ES	EF
A	3	—	0	3
B	4	A	3	7
C	2	A	3	5
D	6	B, C	7	13
E	5	C	5	10
F	3	C	5	8
G	7	E	10	17
H	5	E, F	10	15
I	8	D, G, H	17	25

En este punto hemos determinado realmente la longitud de la ruta crítica. Es el máximo de los primeros tiempos de terminación, o sea 25 semanas en este caso. Sin embargo, para poder identificar las actividades críticas, antes debemos encontrar los tiempos más lejanos.

El cálculo de los tiempos más lejanos se realiza trabajando hacia atrás a través de la red y aplicando esencialmente el dual del procedimiento para los tiempos más próximos. El primer paso sería igualar el tiempo más lejano de todas las actividades que ingresan al nodo final con el valor máximo de los tiempos de terminación más próximos. En este caso hay una sola actividad de terminación, de modo que hacemos

$$LF_I = 25.$$

El tiempo de inicio más lejano se obtiene al restar el tiempo de la actividad, de modo que

$$LS_I = 25 - 8 = 17.$$

Enseguida necesitamos determinar el tiempo de terminación más lejano para todas las actividades que ingresan al nodo 7, que son D, G, y H. Puesto que estas actividades terminan cuando comienza I, tenemos

$$LF_D = LF_G = LF_H = LS_I = 17.$$

Al restar los tiempos de las actividades encontramos los últimos tiempos de inicio para D, G y H. Si resumimos los cálculos hasta este punto, tenemos

Actividad	Tiempo	Predecesor inmediato	ES	EF	LS	LF
A	3	—	0	3		
B	4	A	3	7		
C	2	A	3	5		
D	6	B, C	7	13	11	17
E	5	C	5	10		
F	3	C	5	8		
G	7	E	10	17	10	17
H	5	E, F	10	15	12	17
I	8	D, G, H	17	25	17	25

Debido a que F termina cuando H inicia, $LF_F = LS_H = 12$, y $LS_F = 12 - 3 = 9$. Ahora considere la actividad E. De la figura 9-5, tanto G como H son sucesores inmediatos de E. Esto implica que E debe terminar antes del tiempo de inicio de G y H. Así, el tiempo de terminación más lejano para E es el *primero* de los tiempos de inicio más lejanos de G y H. Es decir, $LF_E = \text{mín}(LS_G, LS_H) = \text{mín}(10, 12) = 10$.

Regla general: El tiempo de terminación más lejano de una actividad es el *mínimo* de los tiempos de inicio más lejanos de sus sucesores inmediatos.

De acuerdo con el diagrama de red de la figura 9-5, tanto E como F son sucesores inmediatos de C. Así $LF_C = \text{mín}\ (LS_E, LS_F) = \text{mín}(5, 9) = 5$, y $LS_C = 5 - 2 = 3$. Dado que B tiene solamente a D como sucesor inmediato, $LF_B = LS_D = 11$, y $LS_B = 11 - 4 = 7$. Finalmente, tanto B como C son sucesores inmediatos de A, de modo que $LF_A = \text{mín}(LS_B, LS_C) = \text{mín}(7, 3) = 3$, y $LS_A = 3 - 3 = 0$.

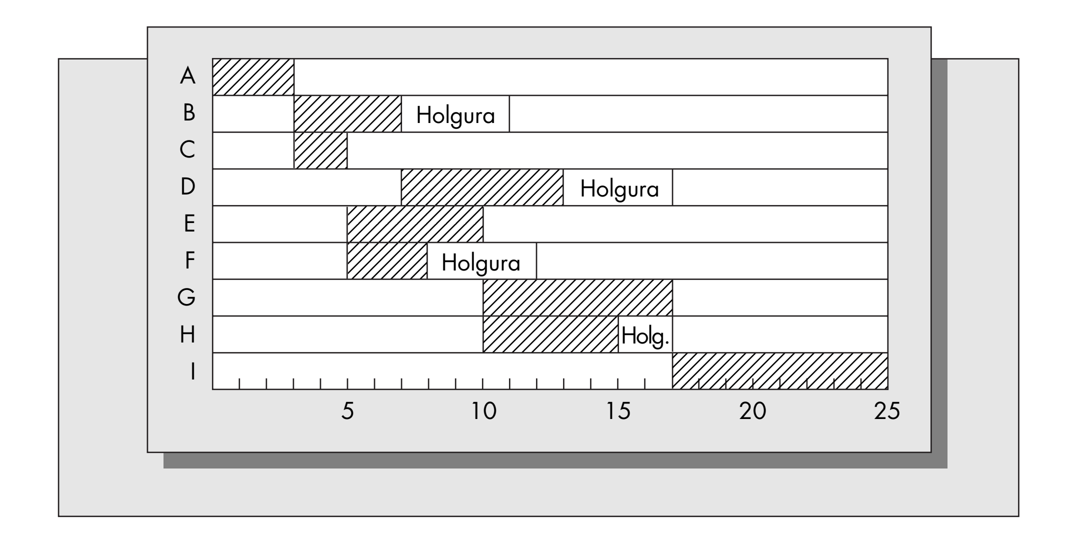
El resumen completo de los cálculos de este ejemplo es

Actividad	Tiempo	Predecesor inmediato	ES	EF	LS	LF	Holgura
A	3	—	0	3	0	3	0
B	4	A	3	7	7	11	4
C	2	A	3	5	3	5	0
D	6	B, C	7	13	11	17	4
E	5	C	5	10	5	10	0
F	3	C	5	8	9	12	4
G	7	E	10	17	10	17	0
H	5	E, F	10	15	12	17	2
I	8	D, G, H	17	25	17	25	0

Hemos añadido una columna rotulada como “holgura”. Ésta es la diferencia entre las columnas LS y ES (también es la diferencia entre las columnas LF y EF). La holgura es el lapso de tiempo en que una actividad se atrasa sin retardar al proyecto. Las actividades con holgura cero —A, C, E, G e I— son las actividades críticas y forman la ruta crítica.

La figura 9-7 es un diagrama de Gantt que muestra los tiempos de inicio y de terminación de cada actividad. Las actividades no críticas se muestran iniciando en los tiempos más próximos, aunque se pueden programar en cualquier momento dentro del lapso marcado por la holgura.

FIGURA 9-7
Diagrama de Gantt para el proyecto de desarrollo de software del ejemplo 9.4.



Ejemplo *9.4 (continuación)*

Suponga que los dos programadores comienzan el proyecto el 1 de junio. Considere las siguientes preguntas:

a) ¿En qué fecha debe iniciar la fusión de los módulos de gráficos y de programas para garantizar que los programadores terminen el proyecto a tiempo?

b) ¿Cuál es el tiempo de terminación más lejano del desarrollo de la pantalla sin retrasar el proyecto?

c) Suponga que North descubre una falla en la codificación del módulo 2, y como consecuencia el tiempo requerido para terminar el módulo 2 es mayor que lo anticipado. ¿Retrasará esto necesariamente el proyecto?

d) Suponga que se descubre un problema similar en el módulo 1. ¿Se retrasará obligatoriamente el proyecto en este caso?

e) Si Bonner es responsable de codificar los módulos y North es responsable del desarrollo de la pantalla, y todas las estimaciones de tiempo son exactas, ¿podrán terminar los programadores el proyecto en 25 semanas?

Solución

a) La fusión de los módulos de gráficos y de programas es la actividad I. Ésta es una actividad crítica, por lo que no debe comenzar después del tiempo de inicio más próximo, que es la semana 17. La fecha de calendario sería el 14 de septiembre.

b) El desarrollo de la pantalla es la actividad D. Ésta no es una actividad crítica y tiene un tiempo de terminación más lejano de 17. La fecha de calendario es nuevamente el 14 de septiembre.

c) El módulo 2 es la actividad F. Como F no es crítica, se permite un retraso de hasta 4 semanas.

d) El módulo 1 es la actividad E. Ésta es una actividad crítica, de modo que una demora retrasaría necesariamente el proyecto.

e) La respuesta es no. El problema es que, suponiendo que Bonner puede trabajar solamente en un solo módulo a la vez, el calendario actual requiere la programación concurrente de los módulos 1 y 2 y de los módulos 3 y 4. Los programadores tendrían que incorporar a un colaborador o subcontratar parte del desarrollo del módulo para terminar el proyecto en 25 semanas.

Problemas para las secciones 9.1 y 9.2

1. Responda a las siguientes preguntas para el ejemplo 9.4.

   a) ¿Cuál evento está representado por el nodo 6 en la figura 9-5?

   b) ¿Cuál grupo de actividades tendrá que terminarse en la semana 16 con objeto de garantizar que el proyecto no se retrase?

c) Suponga que un subcontratista codifica el módulo 4 y lo entrega después de siete semanas. ¿Qué tanto se retrasará el proyecto si el subcontratista no comienza hasta la semana 11?

2. Para el ejemplo 9.4, suponga que los programadores deciden no solicitar ayuda adicional para terminar el proyecto; es decir, Bonner va a codificar los módulos sin ayuda externa.

a) ¿De qué manera alterará esto a la red del proyecto y a las restricciones de precedencia?

b) Encuentre la ruta crítica de la red del proyecto que obtuvo en la parte *a*. ¿Qué tanto es más larga que la determinada en el ejemplo 9.4?

3. Un proyecto que consta de ocho actividades satisface las siguientes restricciones de precedencia:

Actividad	**Tiempo (semanas)**	**Predecesores inmediatos**
A	3	
B	5	A
C	1	A
D	4	B, C
E	3	B
F	3	E, D
G	2	F
H	4	E, D

a) Construya una red para este proyecto. (Deberá necesitar solamente una pseudoactividad.)

b) Calcule los tiempos más próximos y más lejanos de inicio y de terminación de cada actividad e identifique la ruta crítica.

c) Dibuje un diagrama de Gantt del calendario de este proyecto basándose en los primeros tiempos de inicio.

4. Para la red de proyecto ilustrada en la figura 9-8

a) Enliste los predecesores inmediatos de cada actividad.

b) Intente determinar la ruta crítica enumerando todas las trayectorias desde el nodo 1 hasta el nodo 10.

c) Calcule los tiempos más próximos de inicio y de terminación de todas las actividades e identifique las actividades críticas.

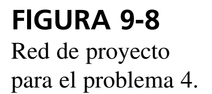


FIGURA 9-8
Red de proyecto
para el problema 4.

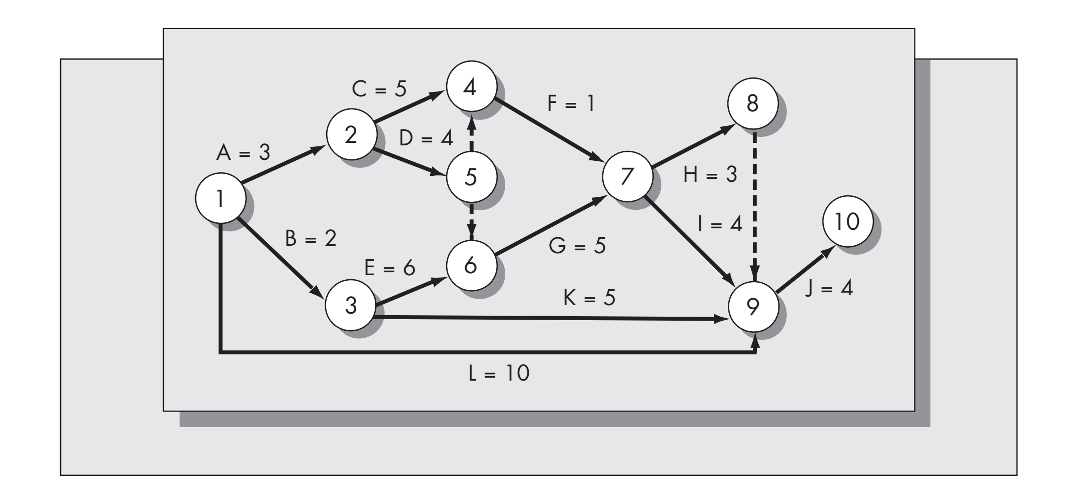
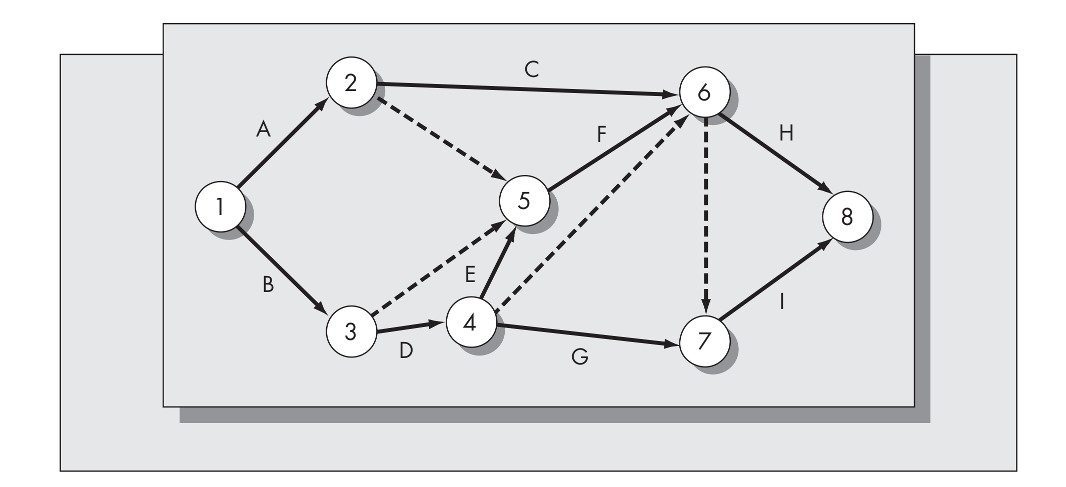


FIGURA 9-9
Red para el problema 5.



5. Considere la red ilustrada en la figura 9-9.

   a) Determine los predecesores inmediatos de cada actividad a partir de la representación de la red. (Sugerencia: asegúrese de considerar solamente los predecesores *inmediatos*.)

   b) Vuelva a dibujar la red basándose en los resultados del inciso *a* con solamente dos pseudoactividades.

6. Un proyecto consiste en siete actividades. Las relaciones de precedencia se dan en la siguiente tabla.

Actividad	Tiempo (días)	Predecesores inmediatos
A	32	
B	21	
C	30	
D	45	A
E	26	A, B
F	28	C
G	20	E, F

   a) Construya una red para este proyecto.

   b) Calcule los tiempos más próximos y más lejanos de inicio y terminación para cada actividad e identifique la ruta crítica.

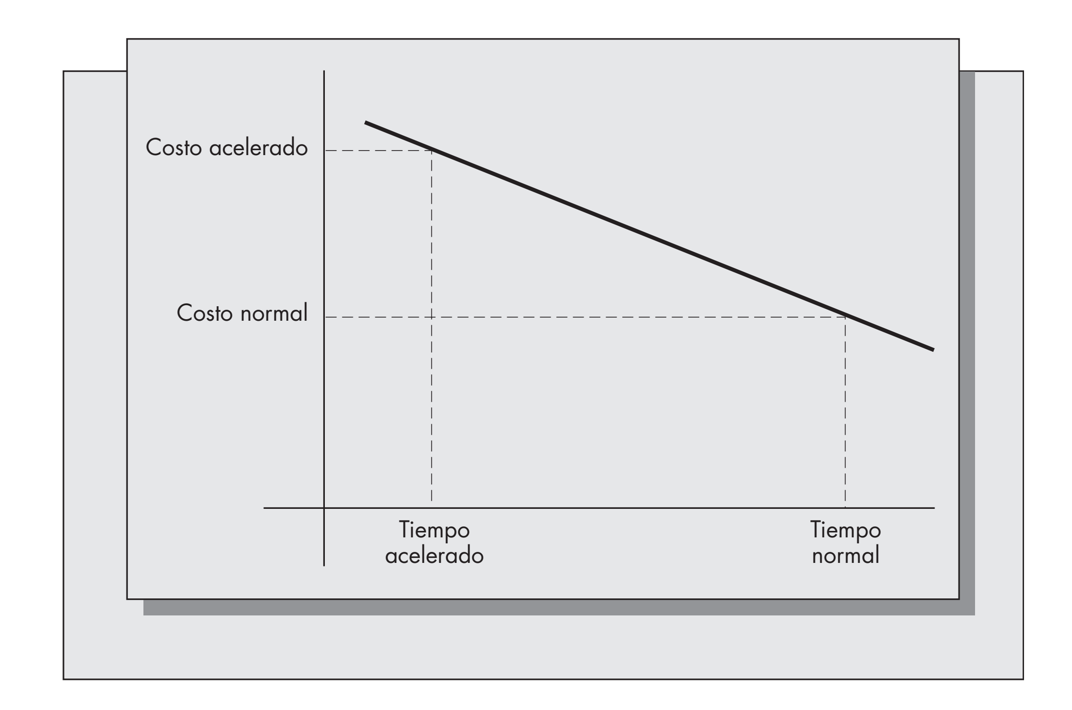
   c) Dibuje un diagrama de Gantt del calendario de este proyecto basándose en los tiempos de inicio más próximos.

   d) ¿Qué grupo de actividades tendrá que terminarse para el día 60 y así garantizar que el proyecto no se retrase?

9.3 MÉTODOS DE COSTEO DE TIEMPO

El CPM ayuda en la programación de proyectos grandes, y además, es una herramienta eficaz para el costeo de los proyectos y para comparar programas alternativos basándose en el costo. En esta sección se analizarán los costos de acelerar las actividades y cómo se incorporan éstos al marco de la administración del proyecto.

FIGURA 9-10
El modelo lineal CPM de tiempo-costo.



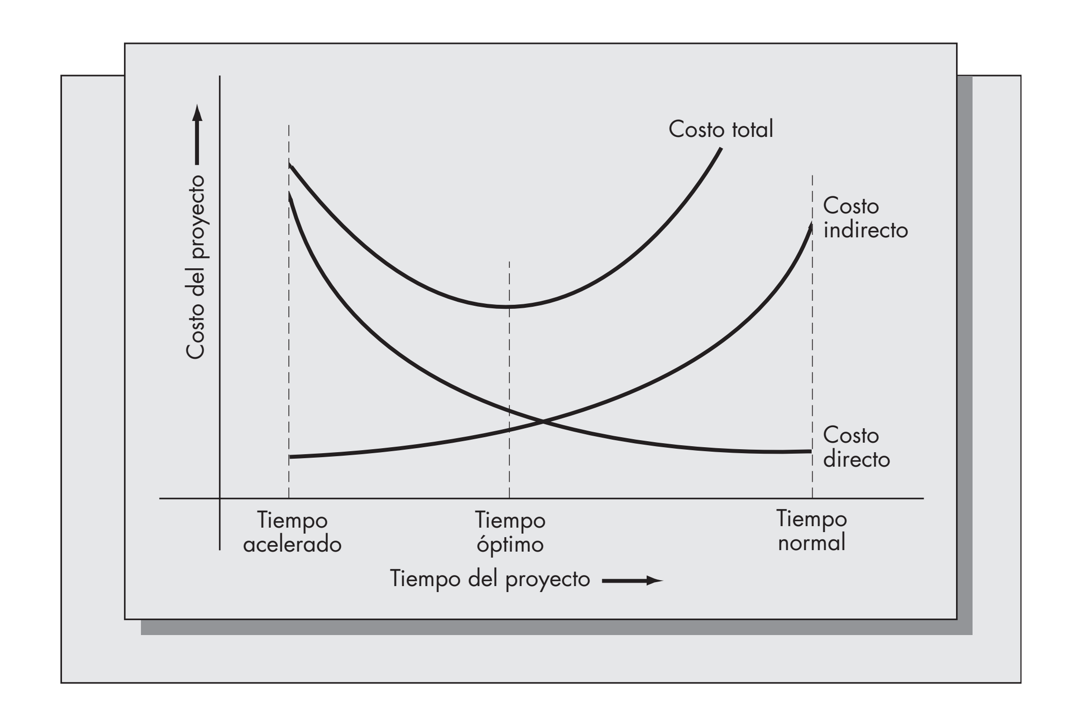
En la sección 9.2 supusimos que el tiempo requerido para terminar cada actividad se conoce y es fijo. Ahora supondremos que los tiempos de las actividades pueden reducirse con un costo adicional. Suponga que el tiempo especificado para terminar una actividad es el *tiempo normal*. El tiempo mínimo posible requerido se define como el *tiempo acelerado*. Además, suponga que se conocen los costos de terminación de la actividad para cada uno de estos instantes de tiempo. Entonces la hipótesis del CPM es que los costos de terminación para los instantes de tiempo entre el tiempo normal y el acelerado están situados a lo largo de una línea recta, como se ilustra en la figura 9-10. En la mayoría de las circunstancias, la hipótesis de que la función del costo de aceleración es lineal deberá ser razonable.

Se cuenta con dos tipos de costos en la mayoría de los proyectos: directos e indirectos. Los primeros incluyen mano de obra, material, equipo, etc. Los segundos incluyen los generales como rentas, intereses, utilidades y cualesquiera otros que aumentan con la duración del proyecto.

Los costos indirectos y los costos directos son respectivamente funciones crecientes y decrecientes del tiempo de terminación del proyecto. Si estas funciones son convexas, la función del costo total, que es la suma de éstas, también será convexa. Esto implica que va a haber un valor para el tiempo del proyecto entre los tiempos normal y acelerado que es óptimo en el sentido de que minimiza el costo total. Las funciones de costo convexas se ilustran en la figura 9-11. La figura 9-10 muestra el costo directo como una función del tiempo para una actividad dada.

El enfoque general es reducir sucesivamente el tiempo del proyecto en una semana (o en cualquier unidad de tiempo para medir las actividades) hasta que ya no sea posible ninguna reducción adicional o hasta que se identifique una solución óptima. Para cada reducción, se calcula el costo directo adicional resultante. Este proceso continúa hasta que se identifique el costo total mínimo. Una dificultad que surge es que a medida que se aceleran las actividades específicas, es posible que las nuevas trayectorias se vuelvan críticas. El procedimiento puede ilustrarse con el uso del ejemplo 9.4 introducido en la sección 9.2.

FIGURA 9-11
Tiempo óptimo de terminación del proyecto.



Ejemplo *9.5*

Retomemos el ejemplo 9.4 relativo a los dos consultores de computación Irving Bonner y Simon North. En la sección 9.2, vimos que es posible que terminen el proyecto en 25 semanas (con alguna ayuda adicional). Una vez que lo hagan y coloquen el programa en el mercado, los consultores esperan recibir un promedio de $1 500 dólares por semana durante los primeros tres años de disponibilidad del producto. Si terminan el proyecto con antelación, esperan recibir este ingreso antes.

Consideran cuidadosamente cada actividad así como la posibilidad de reducir el tiempo invertido y los costos asociados. Los consultores estiman que los costos normales son $500 por semana para actividades que realizan ellos mismos, y más o menos el equivalente para actividades que dan a contrato. La mayoría de las actividades pueden acelerarse subcontratando partes de la programación. De acuerdo con la naturaleza de la tarea los costos de aceleración varían.

Obtienen los siguientes estimados:

Actividad	Tiempo normal (semanas)	Tiempo de aceleración (semanas)	Costo normal	Costo de aceleración	Costo por semana
A	3	1	$ 1 000	$ 3 000	$ 1 000
B	4	3	4 000	6 000	2 000
C	2	2	2 000	2 000	
D	6	4	3 000	6 000	1 500
E	5	4	2 500	3 800	1 300
F	3	2	1 500	3 000	1 500
G	7	4	4 500	8 100	1 200
H	5	4	3 000	3 600	600
I	8	5	8 000	12 800	1 600

La última columna, el costo por semana, muestra la pendiente de la curva de costos que se ilustra en la figura 9-10. Se calcula la pendiente con la fórmula

$$\text{Costo por semana} = \frac{\text{Costo de aceleración} - \text{Costo normal}}{\text{Tiempo normal} - \text{Tiempo de aceleración}}.$$

Se puede calcular el costo total de la ejecución del proyecto en el tiempo normal de 25 semanas al sumar los costos normales para cada una de las actividades. Esta suma da como resultado $29 500. Si se aceleran todas las actividades, el costo total del proyecto aumenta hasta $48 300. ¿Vale este costo agregado los $1 500 por semana de ganancia adicional que los consultores esperan obtener?

Si reemplazamos los tiempos normales por los tiempos acelerados, entonces, al utilizar los métodos de la sección 9.2 (o simplemente por inspección de la red), es fácil demostrar que habrá dos rutas críticas con base en los tiempos acelerados: A–C-E-G–I y A–C–E–H–I. El tiempo de terminación del proyecto se reduce a 16 semanas. El ingreso adicional que obtienen los programadores al reducir el tiempo de terminación del proyecto de 25 a 16 semanas es (9)(1,500) = $13 500, pero el costo adicional de la reducción es $48 300 − $29 500 = $18 800. Así, no es económico reducir todas las actividades a sus tiempos acelerados. (Acelerar todas las actividades implica que tanto las críticas como las no críticas se agilizan. Sin embargo, es evidente que no hay ninguna economía en la aceleración de las no críticas. Entonces, es probable que exista una solución con un tiempo de terminación del proyecto de 16 semanas que cuesta menos de $48 300. El análisis preliminar se usa para obtener una estimación preliminar del costo de reducción del proyecto a su tiempo mínimo.)

Es factible que exista un tiempo de proyecto entre 16 y 25 semanas que sea óptimo. Para determinar el tiempo de proyecto óptimo, encontraremos el incremento del costo directo que resulta de reducciones sucesivas de una semana. Si el incremento del costo es menor a $1 500, entonces se hace evidente que la reducción es económica y deben considerarse reducciones adicionales. Si el costo de la reducción sobrepasa a $1 500, entonces las reducciones adicionales no son económicas y el proceso debe terminarse.

El punto clave a observar es que con objeto de reducir en una semana el tiempo requerido para terminar el proyecto, es necesario reducir el tiempo de una actividad a lo *largo de la ruta crítica presente*, o de actividades a lo largo de las rutas críticas presentes si más de una ruta es crítica. La reducción del tiempo de una actividad no crítica no va a reducir el tiempo del proyecto.

Inicialmente tenemos lo siguiente:

Tiempo del proyecto	Ruta(s) crítica(s)	Actividades críticas	Tiempo actual	Tiempo acelerado	Costo de reducción por una semana
25	A–C–E–G–I	A	3	1	$ 1 000
		C	2	2	
		E	5	4	1 300
		G	7	4	1 200
		I	8	5	1 600

La actividad cuya reducción resulta menos cara es A. Podemos reducir la actividad A a 1 semana sin introducir ninguna nueva ruta crítica. Como el costo de cada reducción semanal es menor que $1 500, es económico reducir A a su tiempo mínimo, que es una semana. En ese punto tenemos lo siguiente:

Tiempo del proyecto	Ruta(s) crítica(s)	Actividades críticas	Tiempo actual	Tiempo acelerado	Costo de reducción por una semana
23	A–C–E–G–I	A	1	1	
		C	2	2	
		E	5	4	$ 1 300
		G	7	4	1 200
		I	8	5	1 600

La siguiente actividad de reducción más barata es G. La ruta crítica va a permanecer igual hasta que G se reduzca a cinco semanas. (Considere la reducción de G en una semana para asegurarse de que ninguna ruta adicional se va a hacer crítica.) Si G se reduce a cinco semanas, ambas rutas A-C-E-G-I y

A-C-E-H-I son críticas. Reducir G de siete a cinco semanas conduce a lo siguiente:

Tiempo del proyecto	Ruta(s) críticas(s)	Actividades críticas	Tiempo actual	Tiempo acelerado	Costo de reducción por una semana
21	A–C–E–G–I				
	A–C–E–H–I	A	1	1	
		C	2	2	
		E	5	4	$1 300
		G	5	4	1 200
		H	5	4	600
		I	8	5	1 600

Con el objetivo de reducir aún más el tiempo del proyecto a 20 semanas, debemos tener la certeza de hacer la reducción a lo largo de ambas rutas críticas. En primer lugar parecería que debemos reducir H porque su costo marginal es el menor. Sin embargo, éste no es el caso. Si reducimos H a cuatro semanas, entonces se reduce solamente la ruta crítica A-C-E-H-I y no la ruta A-C-E-G-I. Si reducimos tanto H como G, el incremento del costo directo es $1 800, y no resulta económico. ¿Significa esto necesariamente que no vale la pena reducir todavía más el tiempo? La respuesta es no.

Observe que las actividades A, C, E e I están situadas simultáneamente a lo largo de *ambas* rutas críticas. Así, una reducción del tiempo de la actividad de cualquiera de esas cuatro actividades va a conducir a una reducción del tiempo del proyecto. Entre estas actividades, E puede reducirse de cinco a cuatro semanas para menos de $1 500. Al hacer esta reducción, obtenemos

Tiempo del proyecto	Ruta(s) críticas(s)	Actividades críticas	Tiempo actual	Tiempo acelerado	Costo de reducción por una semana
20	A–C–E–G–I				
	A–C–E–H–I	A	1	1	
		C	2	2	
		E	4	4	
		G	5	4	$1 200
		H	5	4	600
		I	8	5	1 600

En este punto el costo de reducción del proyecto por una semana adicional sobrepasa los $1 500, de modo que hemos alcanzado la solución óptima. La reducción de 25 semanas a 20 semanas cuesta un total de (1 000)(2) + (1 200)(2) + 1 300 = $5 700 en costos directos adicionales, y conduce a un retorno de (1 500)(5) = $7 500 en ganancia adicional. Si todas las estimaciones de costo y tiempo son correctas, los programadores han obtenido ganancias por $1 800 al considerar el costeo.

Problemas para la sección 9.3

7. Considere el ejemplo 9.5 presentado en esta sección en el cual los dos programadores deben desarrollar un paquete de software comercial.

   a) ¿Cuál es el tiempo mínimo de proyecto y el costo directo total de concluir el proyecto en ese tiempo?

   b) Suponga que todos los tiempos de las actividades se reducen a sus valores mínimos. Explique por qué el costo directo total obtenido en el ejemplo 9.5 es diferente del costo que usted obtuvo en el inciso *a*.

8. Considere el proyecto descrito en el problema 3 con los costos normales y agilizados y los tiempos dados en la siguiente tabla.

Actividad	Predecesor inmediato	Tiempo normal	Tiempo acelerado	Costo normal	Costo acelerado
A	—	3	2	$ 200	$ 250
B	A	5	3	600	850
C	A	1	1	100	100
D	B, C	4	2	650	900
E	B	3	2	450	500
F	E, D	3	2	500	620
G	F	2	1	500	600
H	E, D	4	2	600	900

a) Considere reducciones sucesivas del tiempo del proyecto de una semana y encuentre el costo directo del proyecto después de cada reducción.

b) Suponga que los costos indirectos son de $150 por semana. Encuentre el tiempo óptimo de terminación y el costo total óptimo del proyecto.

9. Discuta la hipótesis de que la curva costo–tiempo es lineal. ¿Qué morfología podría ser más realista en la práctica?

10. Considere el problema 6. Suponga que los costos normales y los acelerados son como se dan en la siguiente tabla.

Actividad	Tiempo normal	Tiempo acelerado	Costo normal	Costo acelerado
A	32	26	$ 200	$ 500
B	21	20	300	375
C	30	30	200	200
D	45	40	500	800
E	26	20	700	1 360
F	28	24	1 000	1 160
G	20	18	400	550

Si los costos indirectos suman $100 diarios, determine el tiempo óptimo de terminación del proyecto y el costo óptimo de terminación del proyecto.

9.4 RESOLUCIÓN DE PROBLEMAS DE RUTA CRÍTICA CON PROGRAMACIÓN LINEAL

Es posible encontrar rutas críticas en las redes de proyectos y también los calendarios de costo mínimo usando, ya sea software reservado para la programación de proyectos o bien programación lineal. Si esperamos resolver los problemas de programación de proyectos sobre una base continua, vale la pena invertir en un producto reservado de software. Sin embargo, la programación lineal es una herramienta útil para resolver un problema de tamaño moderado sobre una base ocasional. Muchos paquetes de programación lineal, como el Solver de Excel, tienen amplia disponibilidad y son fáciles de aprender y de usar. Esta sección muestra cómo formular y resolver los problemas de ruta crítica como programas lineales. (En el resto de este capítulo se supondrá que el lector está familiarizado con la formulación de problemas como programas lineales y con la interpretación de la salida de la computadora. El suplemento 1, que sigue al

capítulo 3, proporciona un análisis de la programación lineal y de cómo pueden resolverse los programas lineales con Excel.)

Con base en la selección de las funciones objetivo, la solución de programación lineal dará como resultado los tiempos de inicio más próximos o más lejanos para cada nodo en la red del proyecto. Primero formulamos el problema del tiempo de inicio más próximo. Suponga que la representación de red del proyecto consiste en los nodos 1 a m, con el nodo 1 que representa el tiempo inicial del proyecto y el nodo m que representa el tiempo de terminación del proyecto.

Sea

x_i = tiempo de inicio más próximo para el nodo i.

t_{ij} = tiempo requerido para terminar la actividad (i, j).[1]

Entonces el tiempo mínimo de terminación del proyecto es la solución del siguiente programa lineal:

$$\text{mín} \sum_{i=1}^{m} x_i$$

sujeta a $x_j - x_i \geq t_{ij}$ para todos los pares de nodos correspondientes a la actividad (i, j).

$$x_i \geq 0 \qquad \text{para } 1 \leq i \leq m.$$

Las restricciones garantizan que hay suficiente tiempo de separación entre los nodos para tomar en cuenta los tiempos de las actividades que representa el arco entre los nodos. A medida que la función objetivo minimiza los valores de x_i, la solución de programación lineal dará como resultado los tiempos de inicio más próximos. Los tiempos de inicio más lejanos pueden encontrarse al reemplazar la función objetivo con

$$\text{mín} \left\{ m x_m - \sum_{i=1}^{m-1} x_i \right\}.$$

Como x_m es el tiempo de terminación del proyecto, todavía deseamos encontrar el valor permisible más pequeño de x_m, de modo que su signo debe permanecer positivo. Ya que cada $x_i \leq x_m$, el multiplicador m para el término x_m garantiza que la función objetivo sea positiva para asegurar que obtengamos una solución acotada. Invertimos el signo de las variables de nodo restantes de modo que la minimización busque sus mayores valores, pero que todavía busque el valor mínimo de x_m. Esto implica que los valores de las variables serán los tiempos de inicio más lejanos para las actividades que emanan de cada uno de los nodos. Una vez que se han determinado los tiempos de inicio más próximo y más lejano para todos los nodos, es fácil traducir esto a los tiempos de inicio más próximo y más lejano para las actividades que usan la representación de red del proyecto. El valor resultante de x_m, el tiempo mínimo de terminación del proyecto, es el mismo para ambas formulaciones.

Ejemplo *9.4 (continuación)*

Para resolver el ejemplo 9.4 usaremos la programación lineal. A partir de la representación de red del proyecto dada en la figura 9-5, vemos que hay un total de ocho nodos. La formulación del problema que da como resultado los tiempos de inicio más próximos es

$$\text{mín} \sum_{i=1}^{8} x_i$$

[1] En las primeras secciones empleamos letras para representar actividades, pero para la mayoría de los sistemas basados en computadora las actividades se representan en la forma (i, j), donde i es el nodo de origen y j es el nodo de destino.

sujeto a lo siguiente:

$x_2 - x_1 \geq 3$	(A)
$x_3 - x_2 \geq 4$	(B)
$x_4 - x_2 \geq 2$	(C)
$x_3 - x_4 \geq 0$	(P_1)
$x_5 - x_4 \geq 5$	(E)
$x_6 - x_4 \geq 3$	(F)
$x_6 - x_5 \geq 0$	(P_2)
$x_7 - x_3 \geq 6$	(D)
$x_7 - x_6 \geq 5$	(H)
$x_7 - x_5 \geq 7$	(G)
$x_8 - x_7 \geq 8$	(I)
$x_i \geq 0$ para $1 \leq i \leq 8$.	

Se muestra entre paréntesis la actividad que da lugar a cada restricción. También deben incluirse las restricciones para las pseudoactividades. Asimismo, la restricción que corresponde a P_1 es $x_3 - x_4 \geq 0$, y no viceversa, ya que P_1 corresponde al arco dirigido del nodo 4 al nodo 3. La parte relevante de la salida de Excel se da aquí.

```
Target Cell (Min)
    Cell        Name        Original Value    Final Value
    $M$7     Min   Value                 0             77

Adjustable Cells
    Cell        Name        Original Value    Final Value
    $B$5      x1                         0              0
    $C$5      x2                         0              3
    $D$5      x3                         0              7
    $E$5      x4                         0              5
    $F$5      x5                         0             10
    $G$5      x6                         0             10
    $H$5      x7                         0             17
    $I$5      x8                         0             25
```

El tiempo mínimo de terminación del proyecto es el valor de x_8, que es 25 semanas. Los tiempos más próximos corresponden más a los nodos que a las actividades, de modo que éstos deben convertirse a tiempos de las actividades. Esto es fácil de hacer si nos remitimos a la representación de red de la figura 9-5. Por ejemplo, como B y C emanan del nodo 2, tendrían tiempos de inicio más próximos iguales a 3 (el valor de x_2). Usted debe sentirse satisfecho de que los tiempos de inicio más próximos para las otras actividades así obtenidos concuerdan con la solución que obtuvimos en la sección 9.2.

Al cambiar la función objetivo a $8x_8 - \Sigma x_i$ y al realizar una nueva corrida con Excel se obtiene la siguiente salida:

```
Target Cell (Min)
    Cell        Name        Original Value    Final Value
    $M$7     Min   Value               148            142

Adjustable Cells
    Cell        Name        Original Value    Final Value
    $B$5      x1                         0              0
    $C$5      x2                         3              3
    $D$5      x3                         7             11
    $E$5      x4                         5              5
    $F$5      x5                        10             10
    $G$5      x6                        10             12
    $H$5      x7                        17             17
    $I$5      x8                        25             25
```

Ahora esto da los tiempos de inicio más lejanos de todas las actividades. Las actividades no críticas son aquéllas con los tiempos de inicio más lejanos que difieren de los primeros; la magnitud de la diferencia es el tiempo de holgura. Debe asegurarse, nuevamente, de que estos resultados concuerden con los obtenidos en la sección 9.2.

Formulación del problema costo-tiempo con programación lineal

Cuando se incluyen los costos de aceleración se puede utilizar la programación lineal para encontrar el tiempo óptimo de terminación. La formulación puede conducir a un programa lineal grande aun para problemas de tamaño moderado. Nuevamente vamos a denominar a las actividades mediante el par (i, j), donde i es el nodo de origen y j es el nodo de destino. Defina a M_{ij} como el tiempo agilizado para la actividad (i, j) y N_{ij} como el tiempo normal para la actividad (i, j). Suponga que la función costo-tiempo (figura 9-10) tiene la representación $a_{ij} - b_{ij}$-t_{ij}, donde a_{ij} es la intersección y, b_{ij} la pendiente y t_{ij} el tiempo de actividad. Sea C el costo indirecto por día. Entonces la formulación de programación lineal del problema que sería encontrar el tiempo óptimo de terminación del proyecto es

$$\text{mín} \sum_{\text{todo}(i,j)} [a_{ij} - b_{ij}t_{ij}] + Cx_m$$

sujeta a

$$\begin{aligned} x_j - x_i &\geq t_{ij} && \text{para todas las actividades } (i, j), \\ t_{ij} &\leq N_{ij}, \\ t_{ij} &\geq M_{ij}, \\ x_i &\geq 0 && \text{para} 1 \leq i \leq m, \\ t_{ij} &\geq 0 && \text{para todas las actividades } (i, j). \end{aligned}$$

Observe que los términos a_{ij} de la función objetivo son constantes y es posible eliminarlas sin alterar la solución. Estas constantes pueden añadirse posteriormente para encontrar el valor óptimo de la función objetivo.

Los tiempos de las actividades, t_{ij}, son ahora variables del problema en vez de constantes dadas. Este problema de programación lineal es considerablemente mayor que el de la sección 9.3. Ahora existe una variable para cada nodo en la red y una variable para cada actividad. Sin embargo, la carga adicional sigue representando menor trabajo que encontrar la solución de costo mínimo manualmente.

Ejemplo *9.5* *(continuación)*

Para resolver el ejemplo 9.5 utilizaremos la programación lineal. Los datos relevantes para el desarrollo de la formulación con programación lineal son

Actividad	Representación de nodo	a_{ij}	b_{ij}	M_{ij}	N_{ij}
A	(1, 2)	4 000	1 000	1	3
B	(2, 3)	12 000	2 000	3	4
C	(2, 4)	—	—	2	2
D	(3, 7)	12 000	1 500	4	6
E	(4, 5)	9 000	1 300	4	5
F	(4, 6)	6 000	1 500	2	3
G	(5, 7)	12 900	1 200	4	7
H	(6, 7)	6 000	600	4	5
I	(7, 8)	20 800	1 600	5	8

El programa lineal resultante es

$$\begin{aligned} \text{Mín}\{&-1\,000t_{12} - 2\,000t_{23} - 1\,500t_{37} - 1\,300t_{45} - 1\,500t_{46} \\ &- 1\,200t_{57} - 600t_{67} - 1\,600t_{78} + 1\,500x_8\} \end{aligned}$$

sujeto a

$$
\begin{aligned}
x_2 - x_1 - t_{12} &\geq 0, \\
x_3 - x_2 - t_{23} &\geq 0, \\
x_4 - x_2 - t_{24} &\geq 0, \\
x_3 - x_4 \qquad &\geq 0, \\
x_5 - x_4 - t_{45} &\geq 0, \\
x_6 - x_4 - t_{46} &\geq 0, \\
x_6 - x_5 \qquad &\geq 0, \\
x_7 - x_3 - t_{37} &\geq 0, \\
x_7 - x_6 - t_{67} &\geq 0, \\
x_7 - x_5 - t_{57} &\geq 0, \\
x_8 - x_7 - t_{78} &\geq 0, \\
1 \leq t_{12} &\leq 3, \\
3 \leq t_{23} &\leq 4, \\
4 \leq t_{45} &\leq 5, \\
2 \leq t_{46} &\leq 3, \\
4 \leq t_{57} &\leq 7, \\
4 \leq t_{67} &\leq 5, \\
5 \leq t_{78} &\leq 8, \\
4 \leq t_{37} &\leq 6, \\
t_{24} &= 2,
\end{aligned}
$$

$$x_i \geq 0 \qquad \text{para } 1 \leq i \leq 8,$$

$$t_{ij} \geq 0 \qquad \text{para todas las actividades } (i, j).$$

Las restricciones de frontera superior e inferior para las variables t_{ij} tendrían que introducirse a la computadora como dos restricciones separadas, dando un total de 28 restricciones.

La salida de Excel para este problema es

```
Target Cell (Min)
     Cell       Name        Original Value    Final Value
     $T$7     Min   Value                0         -19500

Adjustable Cells
     Cell       Name        Original Value    Final Value
     $B$5      x1                        0              0
     $C$5      x2                        0              1
     $D$5      x3                        0              6
     $E$5      x4                        0              3
     $F$5      x5                        0              7
     $G$5      x6                        0              7
     $H$5      x7                        0             12
     $I$5      x8                        0             20
     $J$5      t12                       0              1
     $K$5      t23                       0              4
     $L$5      t24                       0              2
     $M$5      t37                       0              6
     $N$5      t45                       0              4
     $O$5      t46                       0              3
     $P$5      t57                       0              5
     $Q$5      t67                       0              5
     $R$5      t78                       0              8
```

La duración óptima del proyecto está dada por el valor de $x_8 = 20$ y los tiempos óptimos de las actividades son los valores de t_{ij}. Observe que estos tiempos concuerdan con los tiempos de las actividades que encontramos como óptimos en la sección 9.3. El costo total de esta solución se encuentra con la sumatoria Σa_{ij}, que es \$82 700, y el costo de la actividad C, que no se incluyó en la función objetivo. (Podríamos haber incluido el costo de la actividad C en la función objetivo añadiendo el término $1{,}000t_{24}$, ya que t_{24} está fijo como 2.) El valor resultante del costo total del proyecto para 20 días es \$82 700 + \$2 000 − \$19 500 = \$65 200.

Por lo general, la programación lineal no es una manera eficiente de resolver los problemas de red grandes. Los algoritmos que explotan la estructura de red del problema resultan mucho más eficientes. Se dispone de muchos productos de software comercial basados en estos algoritmos para la programación de proyectos. Incluso los productos de software basados en PC tienen capacidad de solución para proyectos grandes. La programación lineal es una herramienta útil para solucionar problemas de tamaño moderado. Sin embargo, para quienes tienen un problema grande de programación de proyectos, o una necesidad continua de programación de proyectos, recomendamos un programa especializado. Véase la sección 9.9 para un estudio exhaustivo del software de administración de proyectos basado en PC.

Problemas para la sección 9.4

11. Resuelva el problema 3 con programación lineal.
12. Resuelva el problema 4 con programación lineal.
13. Formule el problema 5 como un programa lineal.
14. Resuelva el problema 6 con programación lineal.
15. Resuelva el problema 8 con programación lineal.
16. Resuelva el problema 10 con programación lineal.

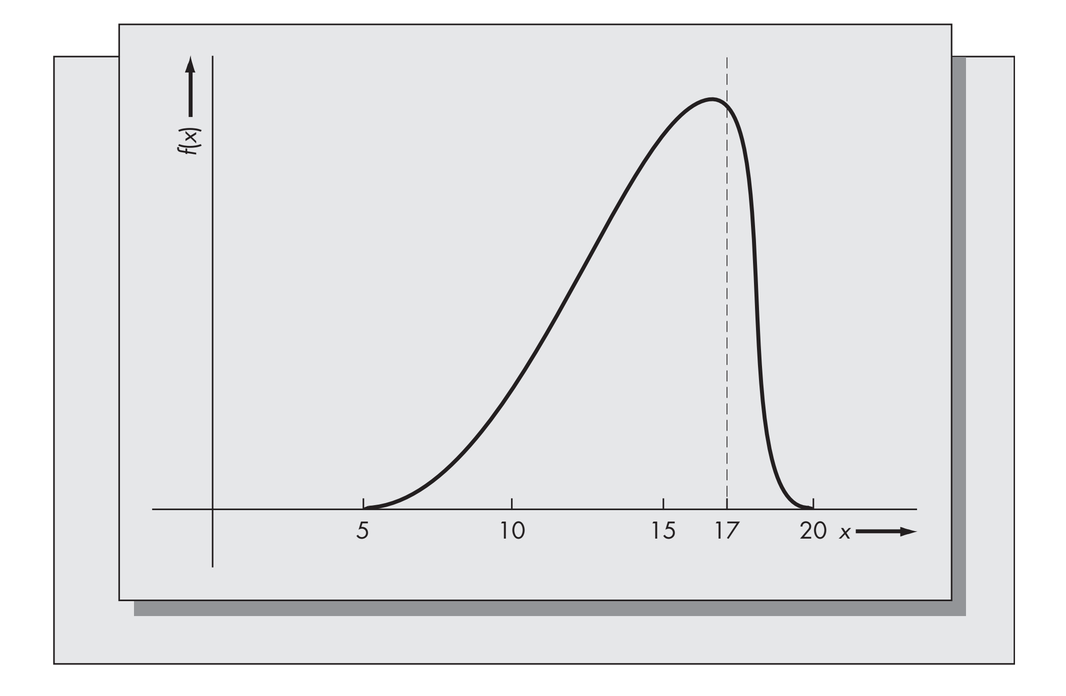
9.5 PERT: TÉCNICA DE EVALUACIÓN Y REVISIÓN DE PROYECTOS

PERT es una generalización del CPM para dar cabida a la incertidumbre en los tiempos de las actividades. Cuando éstos son difíciles de predecir, PERT puede suministrar estimaciones del efecto que la incertidumbre tiene sobre el tiempo de terminación del proyecto. Sin embargo, por razones que se darán con detalle, los resultados del análisis son solamente aproximados. Sea T_i el tiempo requerido para terminar la actividad i. En esta sección supondremos que T_i es una variable aleatoria y que la colección de variables aleatorias $T_1, T_2, \ldots, T_n$ son mutuamente independientes. El primer aspecto que se analizará es la forma apropiada de la distribución de estas variables aleatorias. Ahora defina las siguientes cantidades:

$$a = \text{tiempo mínimo de la actividad},$$
$$b = \text{tiempo máximo de la actividad},$$
$$m = \text{tiempo más probable de la actividad}.$$

Como ejemplo, suponga que $a = 5$ días, $b = 20$ días y $m = 17$ días. Entonces la función de densidad de probabilidad del tiempo de la actividad se ilustra en la figura 9-12. La función de densidad debe ser cero para valores menores que 5 y mayores que 20, y debe ser un máximo para $t = 17$ días. En la teoría de probabilidad, el punto máximo de la densidad se conoce como la *moda*. La *distribución beta* es un tipo de distribución de probabilidad que se define para un intervalo finito que podría tener su valor modal en cualquier punto dentro del intervalo. Por esta razón, generalmente se usa la distribución beta para describir la distribución de los tiempos de las actividades individuales. La hipótesis de los tiempos de las actividades con distribución beta se usa para justificar fórmulas de aproximación simple para la media y la varianza, pero casi nunca se emplea para formular enunciados probabilísticos que se relacionan con las actividades individuales.

FIGURA 9-12
Densidad de probabilidad del tiempo de actividad.



La hipótesis de la distribución beta se usa para justificar las aproximaciones de la media m y la desviación estándar σ para cada tiempo de una actividad. El método tradicional PERT estima μ y σ a partir de a, b, y m usando las siguientes fórmulas:

$$\mu = \frac{a + 4m + b}{6}, \qquad \sigma = \frac{b - a}{6}.$$

La fórmula de la desviación estándar parece basarse en la siguiente propiedad de la distribución normal: los límites a una distancia de 3σ a ambos lados de la media para una variable normal incluyen a toda la población con probabilidad que exceda a .99. En vista de esta propiedad, se supone que hay seis desviaciones estándar desde a hasta b. La fórmula para la media se obtiene suponiendo una aproximación para la varianza así como la distribución beta para el tiempo de la actividad. Dadas las varianzas, a, b, y m, la media se determina mediante una ecuación cúbica. El cálculo de μ para diferentes valores de los otros parámetros y el desarrollo del mejor ajuste mediante regresión lineal conducen al esquema de ponderación de uno-cuatro-uno. (Véase Archibald y Villoria, 1967, p. 449.)

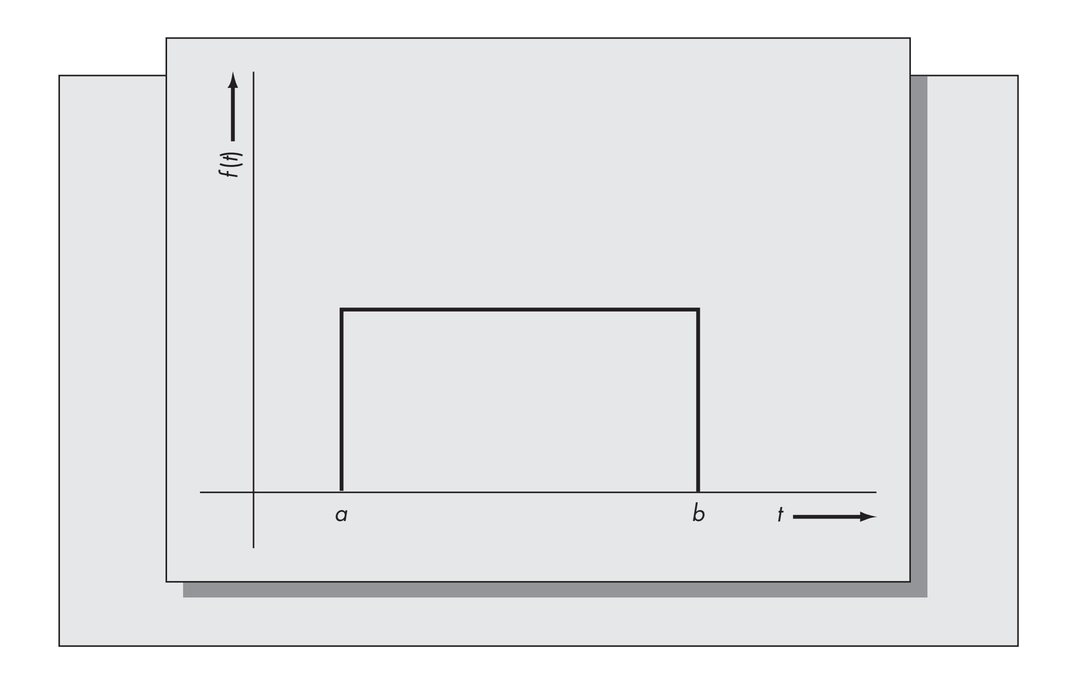
Al elevar al cuadrado ambos lados de la fórmula para σ se obtiene

$$\sigma^2 = \frac{(b - a)^2}{36}$$

donde σ^2 es la varianza.

La distribución uniforme es un caso especial de la distribución beta. Si la duración del trabajo tiene una distribución uniforme desde a hasta b, la función de la densidad sería rectangular, como se ilustra en la figura 9-13. La varianza de una variable uniforme es $(b - a)^2/12$. Debido a que esperaríamos que la varianza fuera menor para una distribución con pico como la que se ilustra en la figura 9-12, la aproximación de σ^2 recomendada anteriormente parece razonable. Observe que el esquema de ponderación uno-cuatro-uno da el valor correcto de la media, $(a + b)/2$, si se sustituye $m = (a + b)/2$ como la moda.

FIGURA 9-13
Densidad uniforme de los tiempos de actividad.



En PERT, se supone que la distribución del tiempo total del proyecto es normal. Para justificar este supuesto se usa el *teorema del límite central*. A grandes rasgos, el teorema del límite central dice que la distribución de la suma de variables aleatorias independientes resulta aproximadamente normal a medida que se incrementa el número de términos de la suma. En la mayoría de los casos, la convergencia se presenta rápidamente. Así que, como el tiempo total del proyecto es la suma de los tiempos de las actividades a lo largo de la ruta crítica, debe tener una distribución aproximadamente normal siempre que los tiempos de las actividades sean independientes.

Suponga que los tiempos de las actividades críticas son $T_1, T_2, \ldots, T_k$. Entonces el tiempo total del proyecto T es

$$T = T_1 + T_2 + \cdots + T_k.$$

Se concluye que el tiempo medio del proyecto, $E(T)$, y la varianza del tiempo del proyecto, $\mathrm{Var}(T)$, están dados por

$$E(T) = \mu_1 + \mu_2 + \cdots + \mu_k,$$
$$\mathrm{Var}(T) = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_k^2.$$

Estas fórmulas encuentran su fundamento en los siguientes hechos de la teoría de probabilidad: el valor esperado de la suma de *cualquier* conjunto de variables aleatorias es la suma de los valores esperados, y la varianza de una suma de variables aleatorias *independientes* es la suma de las varianzas. Se requiere independencia entre los tiempos de las actividades para obtener fácilmente la varianza del tiempo del proyecto y justificar la aplicación del teorema del límite central. Podría modificarse la fórmula de la varianza para calcular correlaciones entre las actividades, pero el supuesto de normalidad del tiempo de terminación del proyecto tal vez ya no sea exacto. Por estas razones, en la práctica no es común el tratamiento explícito de las dependencias entre los tiempos de las actividades.

Resumen del método PERT.

1. Para cada actividad obtenga estimados de a, b, y m. Estos estimados deben ser suministrados por el director del proyecto, o alguien familiarizado con proyectos similares.
2. Con el uso de estos estimados, calcule la media y la varianza de cada uno de los tiempos de las actividades a partir de las fórmulas dadas.

3. Con base en los tiempos medios de las actividades, use los métodos de la sección 9.4 para determinar la ruta crítica.
4. Una vez que se han identificado las actividades críticas, sume las medias y las varianzas de aquéllas para encontrar la media y la varianza del tiempo total del proyecto.
5. Se supone que el tiempo total del proyecto tiene distribución normal con la media y la varianza que se determinan en el paso 4.

Si suponemos que el tiempo del proyecto tiene distribución normal, podremos encarar varios aspectos. Un ejemplo ilustrará el método.

Ejemplo *9.6*

Considere el estudio de caso del ejemplo 9.4 de los dos consultores de computación que desarrollan un proyecto de software. Antes de llevarlo a cabo, deciden que es importante considerar la incertidumbre de los tiempos requeridos para algunas tareas. Al igual que con cualquier proyecto de software, pueden surgir fallas computacionales no esperadas que causen retrasos importantes. A partir de su experiencia, los programadores deciden que los valores de *a*, *b*, y *m* son los siguientes:

Actividad	**Mín (*a*)**	**Más probable (*m*)**	**Máx (*b*)**	$\mu = \dfrac{a + 4m + b}{6}$	$\sigma^2 = \dfrac{(b - a)^2}{36}$
A	2	3	4	3	0.11
B	2	4	10	4.67	1.78
C	2	2	2	2	0
D	4	6	12	6.67	1.78
E	2	5	8	5	1.00
F	2	3	8	3.67	1.00
G	3	7	10	6.83	1.36
H	3	5	9	5.33	1.00
I	5	8	18	9.17	4.69

Hemos calculado las medias y las varianzas de cada uno de los tiempos de actividad con el uso de estos valores. La actividad C, el diseño del diagrama de flujo, requiere de dos días exactamente. Así, la varianza de este tiempo de actividad es cero.

El método tradicional PERT consiste en calcular la ruta crítica en función de los tiempos medios de las actividades. En este ejemplo, la introducción de la incertidumbre no altera a la ruta crítica. Todavía es A-C-E-G-I. (Sin embargo, es posible que en otros casos la ruta crítica basada en los tiempos medios no sea la misma que la basada en los tiempos más probables.) La suma de los tiempos medios de las actividades a lo largo de la ruta crítica, nos proporciona el tiempo esperado de terminación del proyecto, E(T). En este ejemplo,

$$E(T) = 3 + 2 + 5 + 6.83 + 9.17 = 26 \text{ semanas.}$$

En forma similar, la varianza del tiempo de terminación del proyecto, Var(T), es la suma de las varianzas de la actividades a lo largo de la ruta crítica. En este caso,

$$\text{Var}(T) = 0.11 + 0 + 1.0 + 1.36 + 4.69 = 7.16.$$

Se hace el supuesto de que el tiempo de terminación total del proyecto T es una variable aleatoria normal con media $\mu = 26$ y desviación estándar $\sigma = \sqrt{7.16} = 2.68$. Ahora podemos responder a varias preguntas específicas relativas a este proyecto.

Ejemplo *9.7*

Conteste las siguientes preguntas sobre el problema de programación de proyectos descrito en el ejemplo 9.6.

1. ¿Cuál es la probabilidad de que el proyecto pueda terminarse en menos de 22 semanas?
2. ¿Cuál es la probabilidad de que el proyecto requiera más de 28 semanas?
3. Encuentre el número de semanas que se requieren para terminar el proyecto con probabilidad de .90.

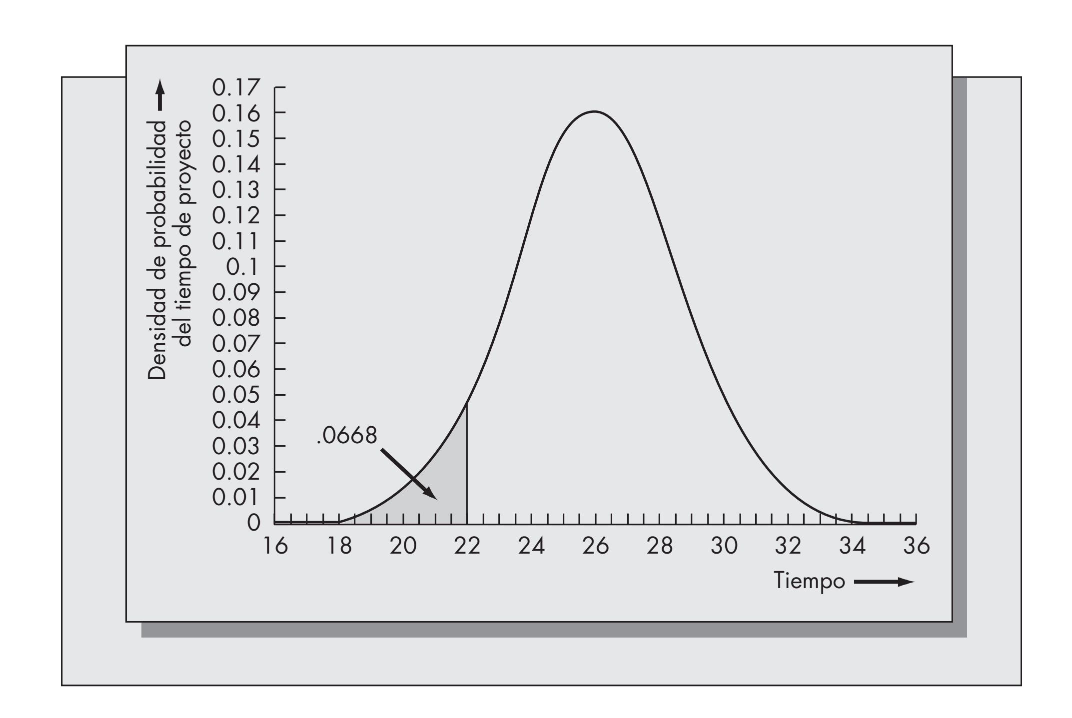


FIGURA 9-14
Respuesta al ejemplo 9.7, parte 1.

Solución

1. Deseamos calcular $P\{T < 22\}$.

$$P\{T < 22\} = P\left\{\frac{T-\mu}{\sigma} < \frac{22-\mu}{\sigma}\right\} = P\left\{Z < \frac{22-26}{2.68}\right\}$$

$$= P\{Z < -1.5\} = .0668$$

Z es la variable normal estándar. La probabilidad proviene de la tabla A-1 de la distribución normal estándar en el apéndice A. La solución se ilustra en la figura 9-14.

2. $P\{T > 28\} = P\left\{\frac{T-\mu}{\sigma} > \frac{28-\mu}{\sigma}\right\} = P\left\{Z > \frac{28-26}{2.68}\right\}$

$$= P\{Z > .75\} = .2266.$$

La solución de este problema se ilustra en la figura 9-15.

3. Aquí queremos encontrar el valor de t tal que $P\{T \le t\} = .90$:

$$.90 = P\{T \le t\} = P\left\{Z < \frac{t-\mu}{\sigma}\right\}.$$

Se concluye que

$$\frac{t-\mu}{\sigma} = z_{.90},$$

o sea $t = \mu + \sigma z_{.90} = 26 + (2.68)(1.28) = 29.43$ semanas.

La notación $z_{.90}$ representa al percentil 90 de la distribución normal estándar. Es decir, $P\{Z \le z_{.90}\} = .90$. Su valor, 1.28, se encuentra en la tabla A-1 del apéndice A en la parte final de este libro. En la figura 9-16 se ilustra la solución.

FIGURA 9-15
Respuesta del ejemplo 9.7, parte 2.

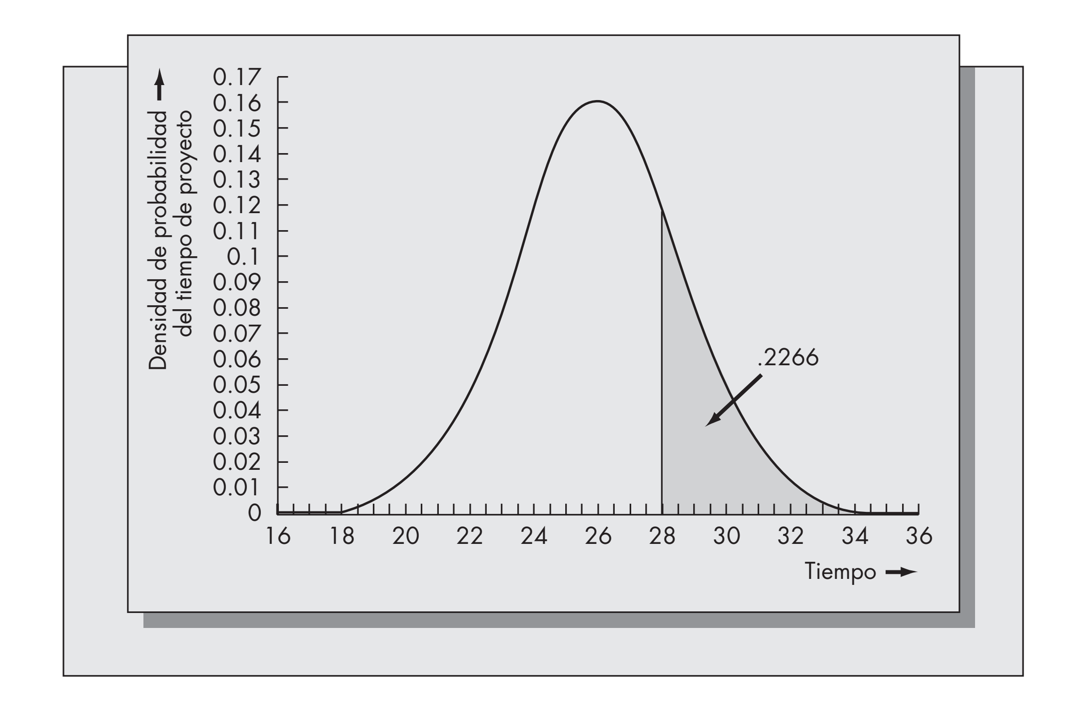
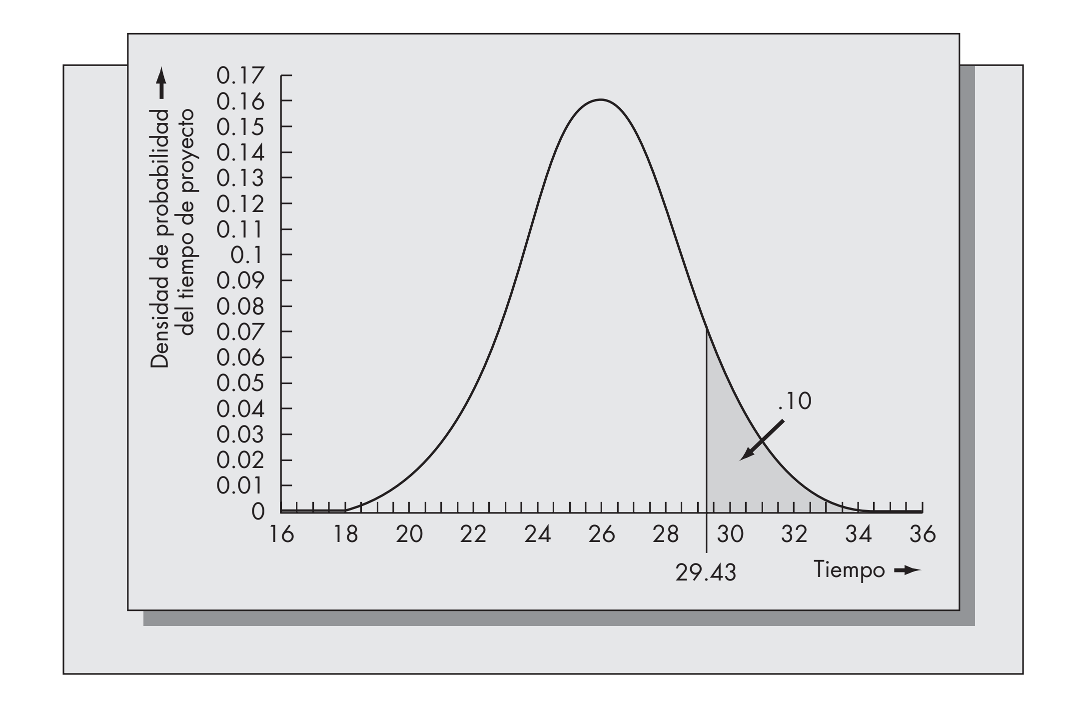


FIGURA 9-16
Respuesta del ejemplo 9.7, parte 3.



Independencia de ruta

Una grave limitación del método PERT es la hipótesis de que la ruta con el tiempo de terminación esperado más largo es necesariamente la ruta crítica. En la mayoría de las redes existe una probabilidad positiva de que una ruta diferente de la que tiene el tiempo de terminación esperado más largo, sea de hecho la crítica. Si éste es el caso, los cálculos de PERT presentados en la sección anterior podrían ser muy engañosos.

Considere el ejemplo 9.7. Suponga que el proyecto concluyó y se observaron los siguientes valores de los tiempos de actividad:

Actividad	Tiempo real requerido para la terminación	Actividad	Tiempo real requerido para la terminación
A	3.4	F	5.0
B	4.0	G	6.2
C	2.0	H	7.2
D	7.0	I	13.0
E	3.5		

Con estos tiempos de actividad, la ruta crítica ya no es A-C-E-G-I, sino que ahora es A-C-F-H-I. Entonces no es exacta la hipótesis de PERT de que una ruta fija es crítica.

Desafortunadamente, determinar la distribución exacta de la ruta crítica es difícil. En general, esta distribución no se conoce. La dificultad es que, como diferentes rutas incluyen las mismas actividades, los tiempos requeridos para terminar las diferentes rutas son variables aleatorias *dependientes*.

El supuesto de la independencia de dos o más rutas puede ser más exacto que el de una sola ruta crítica según la configuración específica de red de proyecto. Observe el siguiente ejemplo.

Ejemplo *9.8*

Considere el proyecto ilustrado en la figura 9-17. En la figura hemos incluido las medias y las varianzas de los tiempos de actividad. Suponga que estos tiempos se expresan en semanas. Hay exactamente dos rutas desde el nodo 1 hasta el nodo 9: A-C-E-G-I y B-D-F-H-I. Los tiempos esperados de estas rutas son 41 y 40 semanas, respectivamente. Con el uso de PERT, supondríamos que la ruta crítica es A-C-E-G-I. Sin embargo, casi siempre existe una posibilidad igual de que la ruta B-D-F-H-I resulte ser crítica tras calcular los tiempos de actividad.

En este ejemplo, las dos rutas tienen solamente a la actividad I en común. Así, éstas son casi independientes estadísticamente. La ejecución de los cálculos suponiendo dos rutas independientes podría dar resultados muy diferentes a partir de la hipótesis de una ruta crítica única.

Suponga que deseamos determinar la probabilidad de que el proyecto concluya en 43 semanas. Sea T_1 el tiempo requerido para terminar la ruta A-C-E-G-I y T_2 el tiempo requerido para terminar la ruta

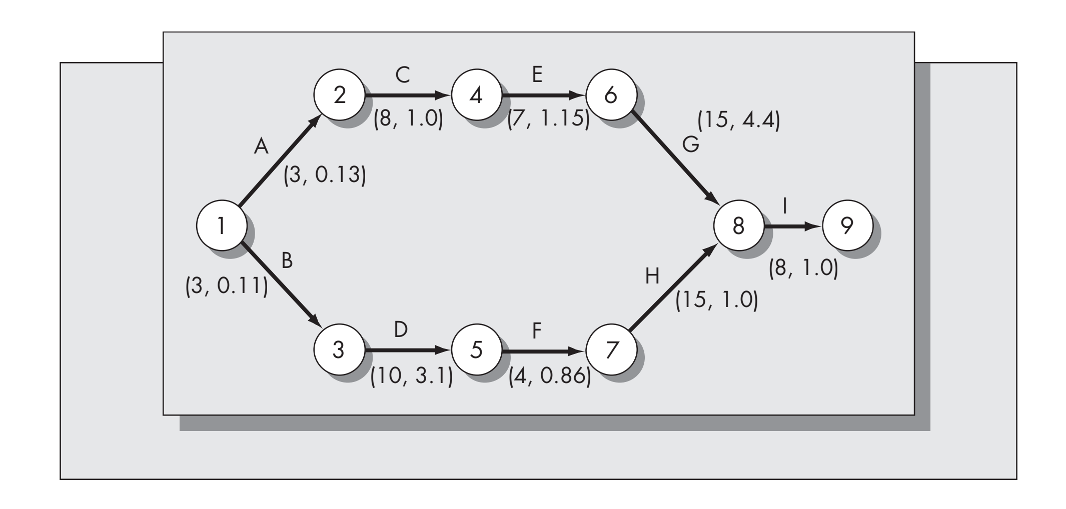


FIGURA 9-17
Red para el ejemplo de PERT 9.8.

B-D-F-H-I. Usando los métodos de la sección 9.4, concluimos que T_1 y T_2 tienen distribución aproximadamente normal, con

$$E(T_1) = 41,$$
$$\text{Var}(T_1) = 0.13 + 1 + 1.15 + 4.4 + 1 = 7.68,$$

y

$$E(T_2) = 40,$$
$$\text{Var}(T_2) = 0.11 + 3.1 + 0.86 + 3.0 + 1 = 8.07.$$

Sea T el tiempo total del proyecto. Es evidente que $T = \text{máx}(T_1, T_2)$. Se concluye que

$$P\{T < 43\} = P\{\text{máx}(T_1, T_2) < 43\} = P\{T_1 < 43, T_2 < 43\}$$

(la última desigualdad se desprende del hecho de que si el máximo de dos cantidades es menor que una constante, entonces ambas cantidades deben ser menores que esa constante)

$$= P\{T_1 < 43\}\, P\{T_2 < 43\}$$

(lo que se concluye de la hipótesis de la independencia de rutas)

$$= P\left\{Z < \frac{43 - 41}{\sqrt{7.68}}\right\} P\left\{Z < \frac{43 - 40}{\sqrt{8.07}}\right\}$$
$$= P\{Z < 0.72\}\, P\{Z < 1.05\}$$
$$= (.7642)(.8413) = .6429.$$

Los métodos de la sección 9.4 habrían dado un estimado de terminación del proyecto en 43 semanas de .7642. Para esta red, .6429 es mucho más exacto.

Ciertos cálculos son más complejos si suponemos independencia de rutas. Por ejemplo, suponga que queremos conocer el número de semanas que se requieren para terminar el proyecto con una probabilidad de .90. Entonces deseamos encontrar t que satisfaga

$$P\{T_1 < t\}\, P\{T_2 < t\} = .90$$

o bien

$$P\left\{Z < \frac{t - 41}{2.77}\right\} P\left\{Z < \frac{t - 40}{2.84}\right\} = .90.$$

El valor de t se calcula por prueba y error. Como es posible que las probabilidades sean cercanas, una buena suposición inicial para el valor de cada una es $\sqrt{.90} \approx .95$, que da $t = 45.6$ para el primer término y $t = 44.7$ para el segundo. El valor correcto es aproximadamente $t = 45.2$ semanas, lo que conduce a un valor de .904 para el producto de las dos probabilidades.

Como las dos rutas de este proyecto tienen solamente una actividad en común, la hipótesis de independencia de rutas es muy razonable y las respuestas obtenidas de esta manera son mucho más exactas que las encontradas suponiendo una ruta crítica única. Sin embargo, en la mayoría de las redes, las rutas pueden tener muchas actividades comunes, y la hipótesis de independencia puede ser inexacta. Considere la red del ejemplo 9.4 que se ilustra en la figura 9-5. En este ejemplo, hay un total de cinco rutas:

A–B–D–I,

A–C–P_1–D–I,

A–C–E–G–I,

A–C–E–P_2–H–I,

A–C–F–H–I.

Las longitudes esperadas de las cinco rutas son, respectivamente, 23.51, 20.84, 26, 23 y 23.17. Desafortunadamente, la ruta A-C-F-H-I contiene tres actividades en común con la ruta esperada más lar-

ga, A-C-E-G-I. En este caso, no resulta claro qué selección proporcionará los resultados más exactos: incluir esta ruta y suponer independencia de rutas o excluirla de toda consideración.

Calcularemos la probabilidad de que el proyecto pueda terminarse en menos de 22 semanas suponiendo independencia de rutas.

Ruta	Tiempo esperado de terminación	Varianza del tiempo de terminación
A–B–D–I	23.5	8.36
A–C–P_1–D–I	20.8	6.58
A–C–E–G–I	26.0	7.16
A–C–E–P_2–H–I	23.0	6.80
A–C–F–H–I	23.2	6.80

Si T es el tiempo de terminación del proyecto y $T_1, \ldots, T_5$ son los tiempos requeridos para terminar cada una de las cinco rutas listadas en la tabla, entonces

$$T = \text{máx}(T_1, \ldots, T_5).$$

Se concluye que

$$\begin{aligned} P\{T < 22\} &= P\{T_1 < 22, T_2 < 22, T_3 < 22, T_4 < 22, T_5 < 22\} \\ &\approx P\{T_1 < 22\}\, P\{T_2 < 22\}\, P\{T_3 < 22\}\, P\{T_4 < 22\}\, P\{T_5 < 22\}. \end{aligned}$$

Nuevamente, suponiendo una distribución normal para cada uno de los tiempos de terminación de ruta, tenemos

$$P\{T_1 < 22\} = P\left\{Z < \frac{22 - 23.5}{\sqrt{8.36}}\right\} = P\{Z < -0.52\} = .3015,$$

$$P\{T_2 < 22\} = P\left\{Z < \frac{22 - 20.8}{\sqrt{6.58}}\right\} = P\{Z < -0.47\} = .6808,$$

$$P\{T_3 < 22\} = .0668 \text{ (del ejemplo 9.7, parte 1).}$$

$$P\{T_4 < 22\} = P\left\{Z < \frac{22 - 23}{\sqrt{6.8}}\right\} = P\{Z < -0.38\} = .3520,$$

$$P\{T_5 < 22\} = P\left\{Z < \frac{22 - 23.2}{\sqrt{6.8}}\right\} = P\{Z < -0.46\} = .3228.$$

Entonces se concluye que

$$P\{T < 22\} \approx (.3015)(.6808)(.0668)(.3520)(.3228) = .0016.$$

El valor verdadero de la probabilidad estará situado entre .0016 y la probabilidad calculada por los métodos tradicionales de PERT, .0668. Se supone que la independencia de rutas puede tener un efecto muy importante sobre las probabilidades. Para este ejemplo, es seguro decir que la posibilidad de que los consultores terminen el proyecto en menos de 22 semanas es mucho menor que 6.68 por ciento y probablemente está muy por debajo de 1 por ciento.

Problemas para la sección 9.5

17. Con referencia al ejemplo 9.4 y a la figura 9-5, ¿cuál es la probabilidad de que el nodo 6 (es decir, la terminación de las actividades A, C, E y F) se alcance antes de que hayan transcurrido 12 semanas?
18. Con referencia al ejemplo 9.4 y a la figura 9-5, ¿cuál es la probabilidad condicional de que el proyecto concluya al final de la semana 25, dado que las actividades de la A a la H (nodo 7) se terminan al final de la semana 15? Para su cálculo, suponga que el tiempo requerido para terminar la actividad I tiene distribución normal.

19. Considere las siguientes estimaciones de tiempo PERT:

Actividad	Predecesores inmediatos	*a*	*m*	*b*
A	—	2	5	9
B	A	1	6	8
C	A	3	5	12
D	B	2	4	12
E	B, C	4	6	8
F	B	6	7	8
G	D, E	1	2	6
H	F, G	4	6	16

a) Dibuje una red para este proyecto y determine la ruta crítica con base en los *tiempos más probables* por inspección.

b) Suponiendo que la ruta crítica es la que usted identificó en el inciso *a*, ¿cuál es la probabilidad de que el proyecto se termine antes de 28 semanas? ¿Antes de 32 semanas?

c) Suponiendo que la ruta crítica es la que usted identificó en el inciso *a*, ¿cuántas semanas se requieren para terminar el proyecto con probabilidad .95?

20. Considere la red del proyecto que se ilustra en la figura 9-18. Suponga que los tiempos asociados a cada nodo son las medias y las varianzas de los tiempos de terminación del proyecto, respectivamente.

a) Identifique todas las rutas de los nodos 1 al 6.

b) ¿Cuál ruta es crítica con base en los tiempos de terminación esperados?

c) Determine la probabilidad de que el proyecto se acabe en 20 semanas suponiendo que la ruta identificada en el inciso *b* sea crítica.

d) Usando solamente la independencia de las rutas A-C-E y B-F-G, recalcule la respuesta para el inciso *c*.

e) Recalcule la respuesta para el inciso *c*, suponiendo la independencia de las rutas identificadas en el inciso *a*.

f) ¿Qué respuesta, *c*, *d*, o *e*, es probablemente la más exacta?

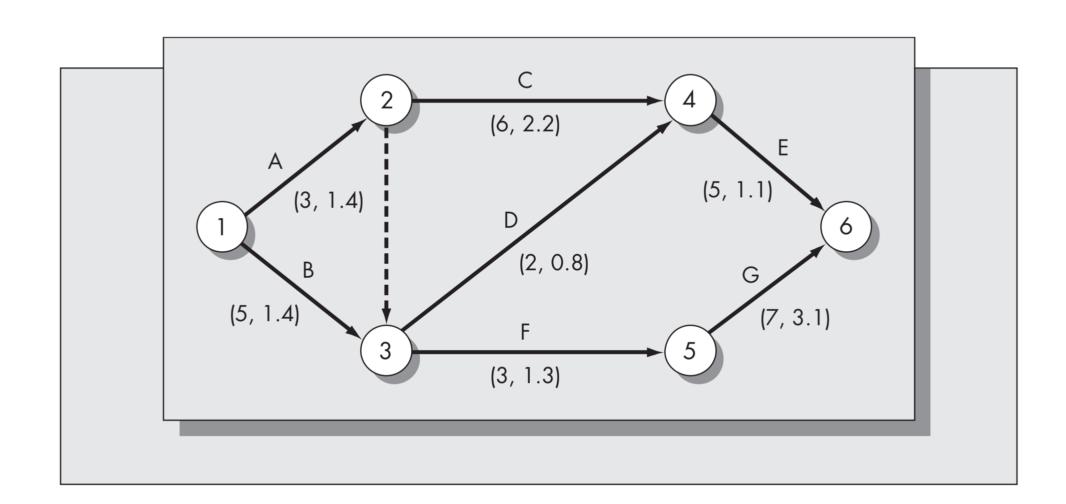


FIGURA 9-18
Red de proyecto (para el problema 20).

21. Considere el proyecto descrito en el problema 3. Suponga que los tiempos de actividad son variables aleatorias con un *coeficiente de variación* constante de 0.2. (El coeficiente de variación es la relación σ/μ.) Suponga que los tiempos dados en el problema 3 son los tiempos medios de actividad.

    a) Calcule la desviación estándar de cada uno de los tiempos de actividad.

    b) Encuentre la media y la varianza de la ruta con el tiempo de terminación esperado más largo.

    c) Usando los resultados del inciso *b*, estime la probabilidad de que el proyecto se termine en 20 semanas.

    d) Usando los resultados del inciso *b*, estime el número de semanas que se requieren para terminar el proyecto con probabilidad .85.

9.6 CONSIDERACIONES DE RECURSOS

Restricciones de recursos para la programación de proyecto único

A lo largo de este capítulo se realiza el supuesto implícito de que se dispone de suficientes recursos y que solamente son importantes las *restricciones tecnológicas* (relaciones de precedencia) para formular los programas. Sin embargo, en la mayoría de los ambientes, las restricciones de los recursos no pueden despreciarse. Los ejemplos de recursos limitados que afectarían a los programas de proyectos son los trabajadores, las materias primas y el equipo. La CPM tradicional no toma en cuenta las consideraciones de los recursos, por lo que un gerente la describió como un "procedimiento factible para producir un programa no factible".

Determinar programas óptimos para redes complejas de proyecto sujetas a limitaciones de recursos es un problema combinatorio muy difícil. Un proyecto individual puede requerir varios recursos diferentes, o muchos proyectos diferentes pueden competir por uno o más recursos. Generalmente se usan métodos heurísticos (aproximados) para modificar los calendarios obtenidos con medios más convencionales.

Un área de gran interés en años recientes ha sido la asignación de recursos escasos entre actividades que compiten. Esta sección considera el proceso de asignación usando el caso del proyecto de programación (ejemplo 9.4) estudiado anteriormente en el capítulo.

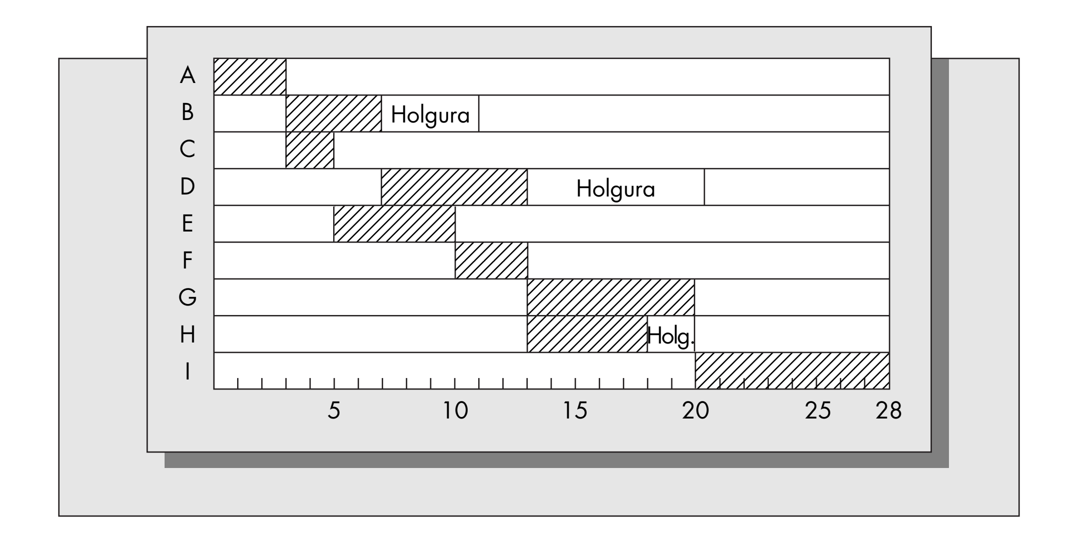
Ejemplo *9.9*

Considere el ejemplo 9.4 acerca de los dos programadores que desarrollan un paquete de software. Un análisis basado solamente en consideraciones de precedencia mostró que el proyecto podría terminarse en 25 semanas sin acelerar ninguna actividad. Sin embargo este análisis se basa en el supuesto de recursos ilimitados. De hecho, como observamos anteriormente, no es posible que los programadores terminen el proyecto en 25 semanas sin ayuda adicional. ¿Cuál es el tiempo mínimo requerido para que terminen el proyecto si no obtienen ayuda adicional?

Suponga que North es responsable de las actividades B (el diseño de los iconos gráficos) y D (el diseño de las pantallas de entrada/salida), y Bonner tiene la pericia necesaria para el desarrollo de los módulos 1 y 2 (las actividades E y F). Además, suponga que cualquiera de los dos programadores puede ejecutar las actividades G y H (los módulos 3 y 4), pero que ambos deben trabajar en la fusión y la prueba final. Considere el diagrama de Gantt ilustrado en la figura 9-7. Éste muestra un programa no factible porque las actividades D, E y H deben ejecutarse simultáneamente entre las semanas 7 y 8, y las actividades D, G y H deben ejecutarse simultáneamente entre las semanas 10 y 13.

Si solamente Bonner y North van a hacer la programación, deben reordenar las actividades sin programar más de dos actividades simultáneamente. Además, deben tener la certeza de que éstas no son dos actividades que solamente uno de ellos pueda ejecutar. Primero, tratan de encontrar un programa factible sin retrasar el proyecto al reprogramar las actividades no críticas dentro de la holgura permisible. De la figura 9-7 es evidente que no importa como se reacomoden las actividades no críticas dentro

FIGURA 9-19
Diagrama de Gantt para el programa modificado del ejemplo 9.9.



de la holgura permisible, no hay manera de evitar la programación de tres actividades simultáneamente en algún punto.

Esto significa que los consultores no pueden terminar el proyecto en 25 semanas sin ayuda adicional. Determinamos un programa factible considerando la secuencia de actividades ejecutadas por cada programador. En la semana 3 North comienza el trabajo en B y Bonner en C. En la semana 5 Bonner comienza el trabajo en E, y en la semana 7 North comienza el trabajo en D. Como F es la codificación del módulo 2, que debe hacer Bonner, ahora F debe seguir a E, de modo que F comience en la semana 10. Tanto G como H pueden comenzar en la semana 13 cuando ambos programadores estén libres. Finalmente, I comienza en la semana 20 y termina en la semana 28. Así, el tiempo del proyecto ha aumentado de 25 semanas a 28 semanas como resultado de las consideraciones de los recursos.

La figura 9-19 muestra el diagrama de Gantt para el programa modificado. Este programa es factible porque nunca hay más de dos actividades programadas para cualquier instante de tiempo. North comienza el trabajo en la semana 3 y trabaja en las actividades B, D, G e I, mientras que Bonner, quien también comienza a trabajar en la semana 3, realiza las actividades C, E, F, H e I. Observe que las actividades G y H son intercambiables.

La inclusión de las consideraciones de los recursos tiene varias consecuencias interesantes. Una de ellas es que la ruta crítica ya no es la misma. Las actividades B, D y F ya no tienen ninguna holgura, así es que ahora son críticas. Además, globalmente hay mucho menos tiempo de holgura.

En general, incluir las restricciones de recursos tiene los siguientes efectos:

1. Se reduce la cantidad total de holgura programada.
2. Puede alterarse la ruta crítica. Además, las actividades con holgura cero no necesariamente van a estar situadas a lo largo de una o más rutas críticas.
3. Los programas con inicios más próximos y más lejanos pueden no ser únicos. Están sujetos a las reglas específicas que se usan para resolver las limitaciones de los recursos.

Existen varios métodos heurísticos para resolver este problema. La mayoría incluye la jerarquización de las actividades de acuerdo con algún criterio y la resolución de los conflictos de los recursos según la secuencia de jerarquización. Algunas de las reglas de jerarquización que se han sugerido incluyen las siguientes:

1. *Holgura de trabajo mínima*. Se da prioridad a las actividades con la menor holgura.

2. *Tiempos de terminación más lejanos*. Cuando exista un conflicto de recursos, esta regla asigna prioridad a la actividad con el mínimo tiempo de terminación más lejano.

3. *Mayor demanda de recursos*. Esta regla asigna prioridad sobre la base de los requerimientos totales de recursos de todo tipo, dando la prioridad más alta a las actividades que tengan la mayor demanda de recursos. El razonamiento que sustenta este método es dar prioridad a las actividades con potenciales cuellos de botella.

4. *Mayor utilización de recursos*. Esta regla da prioridad a aquella combinación de actividades que conduzca a la utilización máxima de los recursos (tiempo muerto mínimo) para cualquier intervalo del calendario.

Davis y Patterson (1975) compararon estas reglas (entre otras). Sus resultados indicaron que los dos primeros métodos de la lista tendían a tener el mejor desempeño.

También existen métodos de solución óptima para la programación de proyectos con restricciones en los recursos. La formulación puede ser un programa entero que requiere algún tipo de procedimientos de rama y límite, o puede resolverse mediante una técnica conocida como enumeración implícita. Para redes grandes estos métodos requieren mucho tiempo de computadora, pero pueden proporcionar soluciones eficientes para redes de tamaño moderado con pocas limitaciones de recursos. Patterson (1984) compara y contrasta tres métodos de solución óptima.

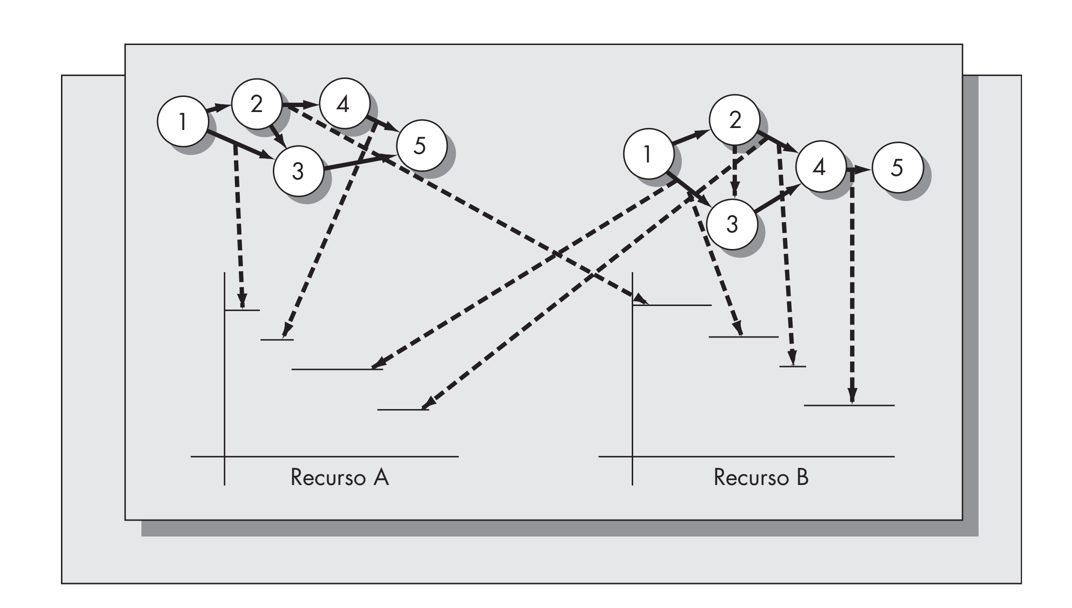
Restricciones de recursos para la programación de proyectos múltiples

Puede resultar complicado el manejo de restricciones de los recursos en los problemas de programación de un solo proyecto, pero las dificultades se magnifican cuando cierto número de proyectos que, además, son independientes comparten un fondo común de recursos. La figura 9-20 muestra este tipo de problema. Dos proyectos requieren los recursos A y B. El retraso de las actividades con objeto de resolver los conflictos de los recursos puede tener consecuencias muy graves para todos los proyectos que requieran el mismo recurso. Existen sistemas computacionales comerciales que tienen la capacidad de manejar decenas de proyectos y tipos de recursos, y posiblemente miles de actividades.

Perfiles de carga de recursos

La planeación de proyectos es útil para programar las actividades interrelacionadas. Un tipo importante del resultado de la planeación incluye los perfiles de carga de los recursos requeridos. Un perfil de carga es una representación con respecto al tiempo de los recursos necesarios. Los perfiles de carga para todos los recursos requeridos se pueden obtener con facilidad en la medida en que se conozcan los requerimientos asociados con cada actividad.

FIGURA 9-20
Dos proyectos que comparten dos recursos.



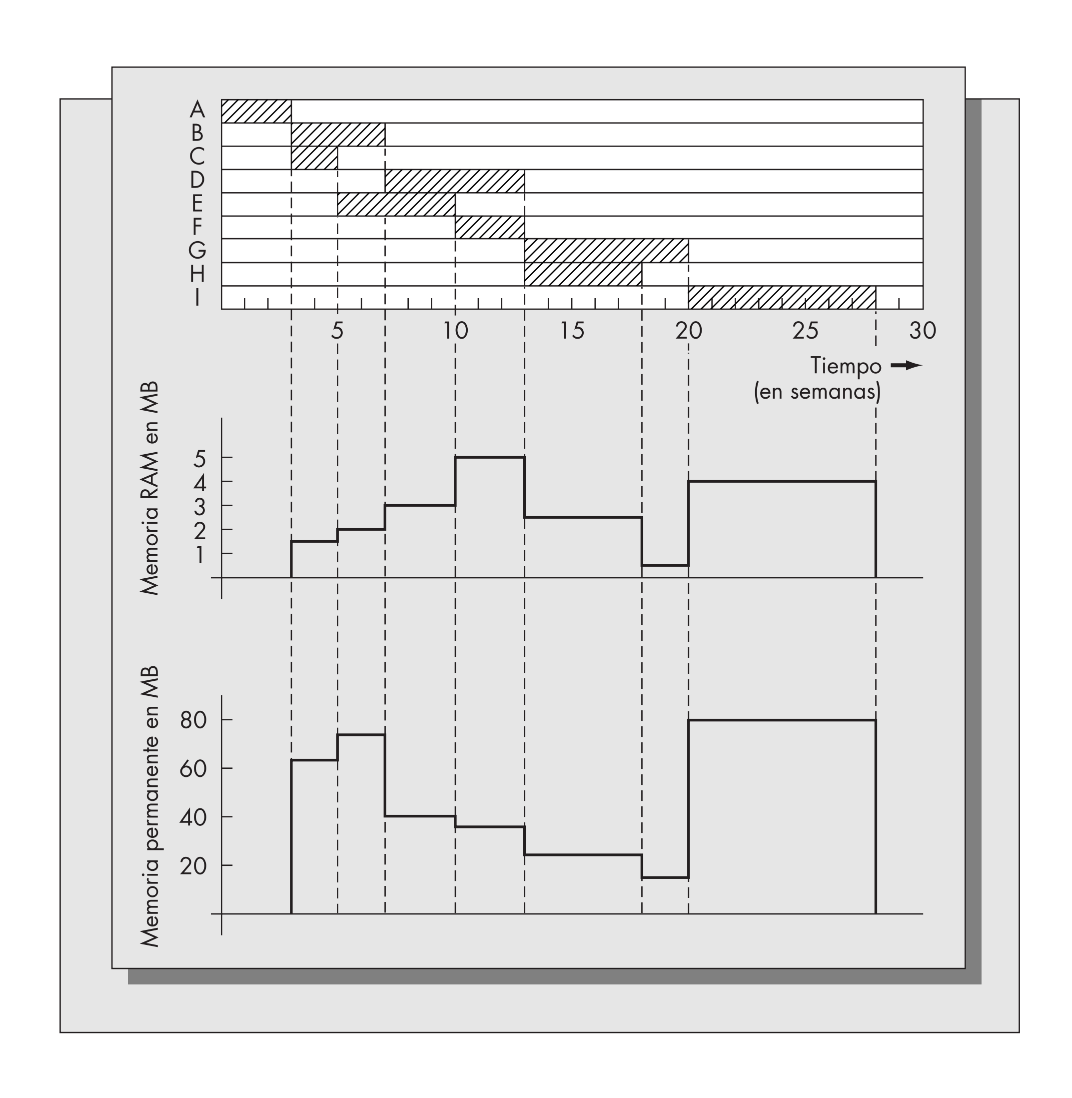
Ejemplo *9.10*

Con el caso de estudio que se introdujo primero en el ejemplo 9.4 ilustramos el procedimiento. Supongamos que los dos programadores están desarrollando el programa en un sistema de cómputo multiusuario individual. El sistema puede segmentar la memoria de acceso aleatorio (RAM) entre los dos usuarios y también puede segmentar el almacenamiento permanente en el disco duro. Tanto la RAM como el almacenamiento permanente se miden en megabytes (MB). Los requerimientos de la memoria RAM y de la memoria permanente para cada una de las actividades que constituyen al proyecto se dan en la siguiente tabla.

Actividad	Requerimiento de RAM (MB)	Requerimiento de almacenaje permanente (MB)
A	0	0
B	1	60
C	0.5	5
D	2	30
E	1	10
F	3	5
G	1.5	15
H	2	10
I	4	80

Suponga que las actividades que constituyen al proyecto siguen los tiempos de inicio y de terminación ilustrados en el diagrama de Gantt de la figura 9-19. La figura 9-21 muestra los perfiles de carga resultantes de la memoria RAM y la memoria permanente. De acuerdo con estos perfiles, el sistema

FIGURA 9-21
Perfiles de carga para la memoria RAM y la memoria permanente (con referencia al ejemplo 9.10).



requerirá cuando menos 5 MB de RAM y 80 MB de almacenamiento permanente. Observe que estamos suponiendo que estos requerimientos no son acumulativos. Los resultados pueden almacenarse en cinta o en disco y pueden recuperarse en un momento posterior cuando haya concluido una parte del proyecto.

Los recursos son consumibles o no consumibles. En el ejemplo 9.10 los recursos eran no consumibles. La fuerza de trabajo es otro ejemplo de un recurso de esta clase. Los recursos consumibles típicos son el dinero en efectivo o los combustibles. Un aspecto que surge con éstos es la cantidad acumulada del recurso consumido para cada instante de tiempo. Exploraremos los perfiles de carga de los recursos consumibles en los problemas al final de esta sección.

Una característica deseable de los perfiles de carga es que sean lo más uniformes posible. Las grandes variaciones de los requerimientos de los recursos dificultan la planeación y pueden conducir a exceder la disponibilidad de los recursos para algún instante de tiempo. La idea que sustenta la *nivelación de recursos* es reprogramar las actividades no críticas dentro de la holgura disponible con objeto de uniformizar el patrón del uso de los recursos. Con frecuencia es posible hacer esta reprogramación por inspección. Para redes más grandes es preferible un método sistemático. Burgess y Killebrew (1962) describen una técnica de nivelación de perfiles de los recursos que se basa en reducir el valor de la suma de los cuadrados de la curva de perfil de recursos reprogramando las actividades dentro de la holgura disponible. No revisaremos el procedimiento aquí.

En resumen, los perfiles de carga de recursos proporcionan un medio importante para determinar los requerimientos impuestos por cualquier programa específico. Dado un programa de proyecto, generalmente se pueden construir los perfiles de carga sin una computadora. Más que un pronóstico estricto de los requerimientos, frecuentemente los perfiles son más útiles como guías preliminares de planeación cuando se esperan variaciones importantes en los tiempos de actividad. De acuerdo con Moder, Phillips y Davis (1983), es probable que se usen más ampliamente que cualquier otra técnica de análisis de recursos.

Problemas para la sección 9.6

22. Para el caso del problema estudiado en este capítulo (ejemplo 9.10), suponga que los requerimientos de RAM y de almacenamiento permanente están dados por

Actividad	Requerimiento de RAM (MB)	Requerimiento de almacenaje permanente (MB)
A	0	0
B	1.5	30
C	2.5	20
D	0.5	10
E	2.0	40
F	1.5	15
G	2.0	25
H	1.5	20
I	4.0	50

a) Determine los perfiles de carga para el almacenamiento RAM y el permanente suponiendo que las actividades se programan de acuerdo con el diagrama de Gantt de la figura 9-7.

b) Determine los perfiles de carga para el almacenamiento RAM y el permanente suponiendo que las actividades se programan de acuerdo con el diagrama de Gantt de la figura 9-19.

23. Considere el proyecto descrito en el problema 3. Se requieren tres máquinas, M1, M2 y M3 para terminar el proyecto. Los requerimientos de cada actividad son los siguientes:

Actividad	Requerimiento de máquinas
A	M1, M2
B	M1, M3
C	M2
D	M1, M2
E	M2
F	M1, M3
G	M2, M3
H	M1

Establezca el tiempo mínimo necesario para finalizar el proyecto si sólo se dispone de una máquina de cada tipo. Al considerar los requerimientos de recursos, ¿cuántas semanas se añaden a la duración del proyecto?

24. Considere el proyecto descrito en el problema 6. Las tareas requieren tanto soldadores como instaladores de tubería. Los requerimientos son

Actividad	Número de soldadores que se requieren	Número de instaladores de tubería que se requieren
A	6	10
B	3	15
C	8	8
D	0	20
E	10	6
F	10	9
G	4	14

a) Determine los perfiles de carga de los soldadores y de los instaladores de tubería suponiendo un programa de inicios más próximos.

b) Determine los perfiles de carga suponiendo un programa de inicios más lejanos.

25. Considere la red de proyecto ilustrada en la figura 9-22. Los tiempos de actividad, medidos en días, se muestran en paréntesis junto a cada etiqueta de actividad.

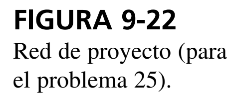
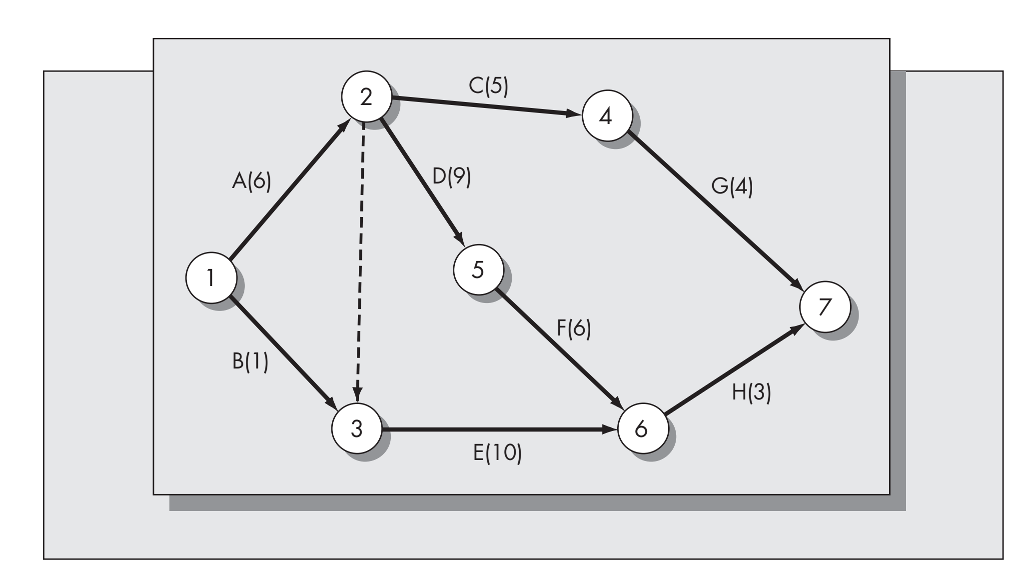


FIGURA 9-22
Red de proyecto (para el problema 25).



a) Establezca los tiempos de inicio y de terminación más próximos y más lejanos de todas las actividades. Dibuje un diagrama de Gantt basado en los tiempos de inicio más próximos, pero indique holgura de actividades cuando sea apropiado. ¿Cuántos días se requieren para terminar el proyecto?

b) Se requiere una sola pieza crítica de equipo para terminar las siguientes actividades: A, B, C, D, G y H. Suponiendo que ninguna de estas actividades puede hacerse simultáneamente, determine un programa factible para el proyecto.

c) Dos recursos, R1 y R2, se usan para cada actividad. Suponga que ambos son recursos consumibles con requerimientos diarios como sigue:

Actividad	Requerimiento diario de R1	Requerimiento diario de R2
A	4	0
B	8	6
C	10	9
D	18	4
E	12	3
F	5	12
G	3	2
H	0	6

Determine los perfiles de carga de recursos a partir del programa encontrado en el inciso *b*.

d) De acuerdo con los resultados del inciso *c*, determine las cantidades *acumuladas* de los recursos R1 y R2 consumidos si se usa el programa encontrado en el inciso *b*.

9.7 CUESTIONES ORGANIZACIONALES EN LA ADMINISTRACIÓN DE PROYECTOS

En este capítulo se han estudiado las técnicas de revisión para ayudar a la función de administración de proyectos. La administración exitosa de proyectos también está sujeta a un efectivo manejo de personal. Un factor importante en el éxito o el fracaso de un proyecto puede ser la forma en que la organización se estructura

La estructura clásica de una organización es la organización en línea. Esto significa que existe una estructura piramidal evidente: los vicepresidentes reportan al presidente, los directores reportan a los vicepresidentes, los gerentes de nivel medio reportan a los directores, etc. En una organización de una línea, generalmente una persona en la base se asigna para coordinar a varios empleados quienes pueden estar en otros departamentos en el mismo nivel de la organización. A la persona responsable se le da el título de líder u oficial de proyecto. La organización de una línea es probablemente la estructura organizativa más débil para administrar proyectos de manera interdepartamental.

La organización divisional por proyectos se encuentra en el otro extremo del espectro organizacional. En este escenario, los empleados están completamente libres con respecto a la organización del y vida del proyecto. Ellos reportan solamente al líder del mismo. Ésta es la estructura organizacional más fuerte para promover la conclusión exitosa del proyecto. Sin embargo, la organización divisional por proyectos requiere que los empleados se roten entre los proyectos; esto puede resultar desconcertante para los trabajadores y costoso para la compañía.

Las compañías han estado experimentando con la organización matricial recientemente. En la organización matricial, la compañía se organiza tanto en sentido horizontal como en sentido vertical. La estructura vertical es la misma que en la organización tradicional en línea. La dirección horizontal corresponde a proyectos individuales que pueden abarcar varios departamentos funcionales. Cada empleado reporta en sentido vertical hacia su superior funcional y en sentido horizontal a su líder de proyecto.

La organización matricial es una mediación entre la organización pura en línea y la organización por proyectos. Sin embargo, cuando los equipos de proyecto abarcan límites departamentales funcionales, surgen problemas. Es común que la lealtad primaria de los empleados se dirija hacia su superior funcional directo. La subordinación dual causa conflictos cuando se formulan exigencias a un empleado desde dos direcciones a la vez. Así, para que el concepto matricial funcione, el líder de proyecto debe tener poder para establecer prioridades y otorgar incentivos al desempeño extraordinario de los miembros del proyecto. Además, se necesita un sentido compartido de las responsabilidades por parte de los miembros del proyecto. Sin embargo, el manejo de lealtades divididas entre los miembros del equipo no es una tarea fácil. Por ejemplo, hace tiempo que la Texas Instruments Corporation adoptó la organización matricial, pero encontró que las líneas ambiguas de autoridad estaban generando problemas.

La estructura del equipo del proyecto varía con la aplicación. Por ejemplo, Vollmann y otros (1992) recomiendan que a un equipo de proyecto al que se le ha asignado la tarea de implementar un sistema de planeación y control de manufacturas consista de cinco a ocho empleados. Lo ideal sería que estuviera compuesto de un representante de cada uno de los departamentos de mercadeo, ingeniería, planificación de producción, manufactura de una línea, y sistemas de información gerencial. También puede quererse un miembro del equipo para las finanzas. Recomiendan que el equipo del proyecto esté liberado de otras responsabilidades y se aísle físicamente de sus departamentos funcionales durante el curso del proyecto (lo que implica que son partidarios de una organización por proyectos). El mejor equipo está constituido por empleados experimentados que provengan de la compañía y que estén familiarizados con los negocios. El líder del equipo de proyecto debe ser un usuario del nuevo sistema, por ejemplo, un profesional en la planeación de la producción o en la manufactura.

9.8 NOTAS HISTÓRICAS

Generalmente se reconoce que los dos sistemas de programación basados en redes, estudiados en este capítulo, CPM y PERT, se desarrollaron casi al mismo tiempo en Estados Unidos a finales de la década de los cincuenta. El CPM fue resultado de un esfuerzo conjunto de la compañía Du Pont y de Remington Rand (Walker y Sayer, 1959). La técnica provino de un estudio dirigido a reducir el tiempo requerido para ejecutar el reacondicionamiento, el mantenimiento y la construcción de la planta. La metodología CPM esbozada en este reporte incluía los intercambios de costo-tiempo de los costos indirectos y directos estudiados en la sección 3, así como los métodos para desarrollar la red del proyecto e identificar la ruta crítica. Al parecer Charnes y Cooper (1962) fueron los primeros en descubrir que la programación lineal podía usarse para resolver los problemas de la programación de proyectos.

PERT se desarrolló como resultado del programa del Sistema de Armas Polaris acometido por la Marina en 1958 (Departamento de Marina, 1958). PERT fue el resultado de un esfuerzo conjunto de la Lockheed Aircraft Corporation, la Oficina de Proyectos Especiales de la Marina, y la compañía consultora de Booz, Allen y Hamilton. Aunque el sistema PERT compartía muchas características con el sistema CPM, el proyecto PERT se concentraba en la incertidumbre de los tiempos de actividad en vez de en los costos del proyecto. Un aspecto interesante relacionado con el enfoque PERT es la justificación de las fórmulas de aproximación para la media y la varianza. Tanto Grubbs (1962) como Sasieni (1986) sacaron a la luz este aspecto.

La fórmula de la varianza se asumió (probablemente con base en las propiedades de la distribución normal) y se obtuvo la fórmula de la media como una consecuencia de la hipótesis de la distribución beta y de la fórmula supuesta para la varianza (véase Archibald y Villoria, 1967, p. 449).

Moder, Phillips y Davis (1983) discuten una referencia escasamente conocida, anterior al desarrollo de PERT y de CPM casi por 30 años, la cual pudo haber sido la verdadera génesis de la metodología de la planeación de proyectos. En 1931 un científico polaco llamado Karol Adamiecki desarrolló y publicó un dispositivo de planeación conocido como la gráfica de armonía (Adamiecki, 1931). Ésta consiste, básicamente, en un diagrama vertical de Gantt modificado para incluir predecesores inmediatos. La técnica requiere usar barras deslizantes para cada actividad. (Véase Moder, Phillips y Davis, 1983, pp. 10-12, para una discusión más exhaustiva del método de la gráfica de armonía.)

La investigación de las redes de proyectos continúa. Un área de constante interés son las redes de recursos restringidos (Patterson, 1984). Otra son las redes con tiempos de actividad aleatorios. Como lo indicamos en la sección 9.5, la metodología PERT proporciona sólo una aproximación de la distribución del tiempo en que el proyecto se termina. El interés se ha centrado recientemente en el desarrollo de simulaciones eficientes (Sullivan, Hayya y Schaul, 1982) o en determinar la distribución exacta del tiempo de conclusión del proyecto suponiendo una distribución diferente a la beta para el tiempo requerido por cada actividad (Kulkarni y Adlakha, 1986).

9.9 SOFTWARE DE ADMINISTRACIÓN DE PROYECTOS PARA PC

A medida que aumenta la base instalada de computadoras personales, la demanda de software también aumenta. El software de administración de proyectos estuvo disponible poco después que las computadoras mainframe se vendieron a los negocios. La mayoría de los primeros fabricantes de computadoras (IBM, Honeywell, Control Data, entre otros) comercializaron algún tipo de software de administración de proyectos en una época tan temprana como la década de los cincuenta. Con la difusión de las computadoras personales en la comunidad de negocios en los ochenta, los distribuidores de software se dieron cuenta de que había un mercado importante para el software de administración de proyectos basado en PC. Las herramientas de administración de proyectos, junto con los procesadores de palabras y las hojas de cálculo, se convirtieron en componentes importantes del menaje de software disponible para las computadoras personales.

En la edición anterior de este libro se revisaron varios programas de administración de proyectos. La mayoría (pero no todos) de estos productos todavía están disponibles. Por ejemplo, todavía es posible hallar el programador de proyectos de Scitor para Windows (actualmente se cuenta con la versión 8.5, llamada frecuentemente PS Suite). Ahora está disponible el CA-Superproject en la versión 4.0, que todavía comercializa Computer Associates. Aun cuando Symantec sigue activo, su paquete de software de administración de proyectos, Time Line, ya no está disponible. Como se esperaría, Microsoft sigue dominando esta área. Microsoft Project para Windows se introdujo originalmente en 1990 (para reemplazar versiones anteriores de MSDOS). Se diseñó como un producto a nivel de introducción, y no ofrecía muchas de las características sofisticadas que los profesionales de la administración de proyectos necesitan. Actualmente, Microsoft es un proveedor líder del software de administración de proyectos. Microsoft todavía vende y apoya a Microsoft Project en el mercado masivo. Las salas de chateo y los tableros de boletín para los usuarios son comunes en el MS Project. Además, Microsoft ofrece tres productos compatibles con la administración de proyectos diseñados para el mercado profesional. Éstos son Microsoft Project Server, Microsoft Project Professional y Microsoft

Project Web Access. Estos productos trabajan juntos y se integran sin problemas en el Microsoft Office Suite. Proporcionan al usuario las características y la sofisticación que requieren los profesionales de la administración de proyectos, incluyendo la administración de proyectos múltiples, la administración de recursos, y el manejo sofisticado de los datos.

El software basado en la Red es uno de los desarrollos más recientes para la administración de proyectos. En el periodo de 1998-2000 parecía que la Red se apoderaría del mundo. Una compañía que mencionara un sitio electrónico en su literatura promocional veía que el precio de sus acciones se disparaba al cielo con base solamente en ese anuncio. Se estimaba que para que Amazon.com generara suficientes utilidades con miras a justificar el alza en sus acciones durante 2000, habría tenido que vender libros a todo el mundo. La capitalización de mercado de Cisco Corporation, un fabricante de buriladoras con su matriz en el norte de California, sobrepasó a la de compañías de diez veces su tamaño o más grandes tales como IBM, General Motors y AT&T. De hecho, en su punto más alto, Cisco tuvo la capitalización de mercado más grande de *cualquier* compañía estadounidense. Sin embargo, al igual que la desenfrenada especulación con los bulbos de tulipán en Holanda en el siglo XVII llegó a un abrupto fin, la especulación desbordante con las acciones de Internet corrió la misma suerte. Algunas personas en todo el país se hicieron millonarias en un instante, cuando subieron hasta el cielo los precios de las acciones de sus lanzamientos, y lo perdieron todo igual de rápido cuando estos precios se colapsaron. De hecho, la gran mayoría de los lanzamientos de la Internet (incluyendo muchas compañías razonablemente exitosas como Webvan y Pets.com) quebraron.

Pero la quiebra del dot-com de 2001 no implicó el fin de la creciente influencia de la Internet. Su uso siguió incrementándose y muchas de las compañías que lograron quedar en pie cuando se disipó el polvo llegaron a ser bastante exitosas. Más y más aplicaciones de software, que antes estaban disponibles sólo para operarse independientemente, migraron a la Red. Lo mismo ocurrió con el software de administración de proyectos. Por ejemplo, en los inicios de 2004 el sitio electrónico http://www.comp.glam.ac.uk/pages/staff/dwfarthi/projman.htm#sw listaba 18 paquetes de software diferentes de administración de proyectos basados en la Red. Uno, el AceProject, está disponible gratuitamente, y puede instalarse en la Intranet de una compañía. Otro paquete bajado de la Red, ProjectCompanion, tiene la ventaja adicional de funcionar tanto en la Red como en PDA basados en Palm manuales. Otros paquetes, tales como el Xcolla, están dirigidos al suministro de una plataforma para colaboración por equipos y rastreo de tareas. La ventaja de un sistema basado en la Red es evidente. Facilita la colaboración en la función de administración de proyectos entre partes múltiples. Los trabajadores en diferentes instalaciones y/o localidades en todo el mundo tienen acceso en tiempo real a los mismos archivos y a la misma información. El que un paquete de administración de proyectos basado en la Red pueda usarse en una PDA es una bendición de lo portátil. El número de sitios con acceso inalámbrico a Internet se incrementa diariamente, y se han sumado campus universitarios, aeropuertos y hasta Starbucks. Esto significa que los profesionales tienen acceso a archivos compartidos en cualquier parte del mundo en dispositivos que pesan unas cuantas onzas y que tienen baterías con una duración de días en lugar de horas.

Observe que todos los productos de software mencionados en esta sección son herramientas de administración de proyectos de uso general. Sin embargo, hay muchos paquetes de administración de proyectos que están reservados a aplicaciones específicas. Varios están disponibles para ayudar al departamento jurídico a manejar sus cargas de trabajo. Se dispone de paquetes basados en la Red y que operan independientemente de la computadora para las aplicaciones de construcción y de edificación. Éstos son sólo dos ejemplos. La demanda de software de administración de proyectos para usos especiales y para uso general sigue creciendo.

Aplicación instantánea

LA ADMINISTRACIÓN DE PROYECTOS AYUDA A UNITED A CUMPLIR CON EL PROGRAMA

United Airlines ha sido usuario por mucho tiempo de los métodos y del software de administración de proyectos lo que les ha posibilitado mantenerse a flote. United posee 549 aviones y realiza 2 000 vuelos cada día. Para que una operación de esta magnitud funcione apropiadamente se necesitan dedicación por parte de los empleados y herramientas de calidad para la administración de proyectos. Richard Gleason, gerente de la oficina de proyectos MIS en United, eligió Project Workbench, un paquete de software basado en PC de la corporación Advanced Business Technology en la ciudad de Nueva York, para llevar a cabo las actividades..

United usa este paquete para integrar más de 80 aplicaciones de mainframe en distintas áreas como el desarrollo arquitectónico y el soporte técnico para una fecha de entrega operativa. Además, el software ayuda a United a administrar su programa de mantenimiento de la flota a nivel mundial.

El software permite que United enlace dependencias entre un número ilimitado de proyectos diferentes y calcule calendarios para estos proyectos. El software es más caro que la mayoría de uso general dirigido al mercado de masas (véase la discusión sobre el software de administración de proyectos en la sección 9.9). Sin embargo, United piensa que el costo adicional queda más que compensado por la flexibilidad de programar proyectos múltiples con enlaces de actividades dependientes (Ouelette, 1994).

THOMAS BROTHERS PLANEA LA CONTRATACIÓN DE PERSONAL CON SOFTWARE DE ADMINISTRACIÓN DE PROYECTOS

Thomas Brothers Maps, Inc., con base en Irvine, California, es uno de los más destacados editores de mapas de carreteras del país. La compañía ofrece guías viales y mapas que cubren muchas ciudades y condados de la costa oeste. Tiene 230 trabajadores y ganancias anuales que sobrepasan los $20 millones. Bob Foster, el presidente de la compañía, comenzó a usar software de administración de proyectos cliente-servidor con el fin de planear un programa para contratar y entrenar cartógrafos. Foster escogió PlanView, un paquete basado en Windows, de la compañía homónima con sede en Austin, Texas. El software maneja la planeación de los 250 proyectos que la compañía acometió en 1995. El paquete de administración de proyectos permite al presidente rastrear con exactitud el despliegue de cada trabajador.

Una de las características de PlanView que atrajo a Foster y a otros empleados a Thomas Brothers fue la capacidad del software de rastrear recursos así como calendarios. PlanView contempla la carga de trabajo en términos de sobrecarga de recursos más bien que como una sucesión de tareas y cuellos de botella potenciales. La compañía puede rastrear su progreso en un gran número de proyectos múltiples al mantener un programa en función de la carga de trabajo para cada cartógrafo. De acuerdo con Foster, "nunca habíamos podido ver tan lejos con tanta claridad".

Otra característica de PlanView es que el paquete usa una base de datos SQL estándar y fácilmente puede establecer una interfase con otros productos de bases de datos (Oracle 7 en este caso). La información de la administración de proyectos suministrada por el sistema puede usarse para rastrear también otras funciones de los negocios, como el rastreo del costo de las ventas (Word, 1995).

FLORIDA POWER AND LIGHT TOMA EN SERIO LA ADMINISTRACIÓN DE PROYECTOS

Florida Power and Light (FPL), una importante compañía de servicios, maneja dos instalaciones de generación de energía nuclear en Turkey Point y St. Lucie. FPL estableció un departamento independiente de administración de proyectos para manejar la administración de estas instalaciones, así como otras varias funciones dentro de la compañía. En su apogeo, el departamento tenía 17 empleados de control de proyectos para apoyar a más de 600 ingenieros y analistas. Se establecieron grupos separados especializados para cada planta.

El grupo estableció requerimientos de control de proyectos y de reportes de ejecución para todo trabajo importante contratado con el objetivo de negociar eficazmente con los contratistas. Este proceso requirió que FPL identificara a sus principales contratistas y que iniciara negociaciones con los mismos para implementar el sistema. Resultó que cuatro contratistas eran responsables de 80 por ciento de la carga de trabajo. Cada uno de estos contratistas había implementado un sistema satisfactorio de control de proyectos en menos de un año.

Para hacer que el sistema fuera más amigable con el usuario, FPL abandonó su software de *mainframe* tradicional y lo reemplazó con el software de Welcom basado en PC. Aun cuando el sistema basado en PC no era tan poderoso como el paquete de mainframe para manejar los aspectos del modelado de los recursos, la impresión era que la propiedad local del proceso brindado por las PC compensaría las limitaciones del nuevo software.

La autoevaluación de la función de administración de proyectos de FPL mostró que los clientes estaban muy satisfechos con la función de administración de proyectos y con el soporte de programación ofrecido por el grupo de administración de proyectos. De hecho, a FPL se le concedió el premio Deming hace algunos años por su dedicación a la mejora continua y a la calidad del servicio. El éxito de sus iniciativas de administración de proyectos jugó un papel importante en este logro (Cooprider, 1994).

9.10 Resumen

Las técnicas de administración de proyectos proporcionan útiles herramientas para preparar y administrar los programas de proyectos grandes que constan de muchas actividades interrelacionadas y que deben terminarse en una secuencia específica. La planeación se puede aplicar en muchos niveles de una organización, para proyectos que duran desde unos cuantos días hasta meses o aun años.

Este capítulo se centró en el método de la ruta crítica (CPM) y sus ampliaciones. Las *redes* son una manera conveniente de representar un proyecto. Se cuenta con dos maneras de usar las redes para representar a los proyectos: actividades en las flechas y actividades en los nodos. Al usar el método de actividades en las flechas, los nodos de la red corresponden a la terminación de algún subconjunto de actividades. Si se usan nodos en vez de flechas para representar a las actividades, no se requieren pseudoactividades. Sin embargo, las actividades en las flechas se usa mucho más en la práctica y en general es más intuitiva.

La *ruta crítica* es la ruta o cadena *más larga* a través de la red. Su longitud es el tiempo mínimo de terminación del proyecto y las actividades que se sitúan a lo largo de ésta se conocen como críticas. Una demora en una actividad crítica retrasa al proyecto. Las actividades no críticas tienen *holguras*, lo que significa que pueden atrasarse sin retrasar necesariamente al proyecto. Aunque para redes pequeñas la ruta crítica puede identificarse fácilmente por inspección, para redes más grandes se requiere un procedimiento algorítmico. Este capítulo presentó un método que incluye tanto pasos hacia delante como hacia atrás a través de la red, y especifica los tiempos de inicio y de terminación más próximos y más lejanos para todas las actividades.

Una de las metas del trabajo inicial de desarrollo en CPM era considerar el efecto del *costeo de proyecto*. Suponemos que los costos son directos o indirectos. Los primeros incluyen la mano de obra, el material y el equipo; éstos aumentan si disminuye la duración del proyecto. Los indirectos incluyen los costos de las rentas, el interés y las utilidades; éstos se incrementan si la duración del proyecto aumenta. El objetivo del análisis radica en determinar el tiempo óptimo de ejecución del proyecto que minimice la suma de ambos tipos de costos.

La *programación lineal* es uno de los medios para solucionar problemas de programación de proyectos. Las formulaciones de programación lineal consideradas aquí resuelven tanto el problema de CPM como el problema de CPM con intercambios de costo-tiempo. Aunque no se estudia en este capítulo, la programación lineal también puede usarse para resolver algunos problemas de costo-tiempo si los costos directos son funciones no lineales de los tiempos de duración de la actividad.

PERT es una ampliación del análisis de ruta crítica para incorporar a la incertidumbre en los tiempos de actividad. Para cada actividad, se requieren tres estimados de tiempo: (1) el tiempo mínimo de actividad (llamado a), (2) el tiempo máximo de actividad (llamado b) y (3) el tiempo más probable de actividad (llamado m). A partir de tales estimados, se aproximan la media y la desviación estándar del tiempo de actividad. Se supone que la duración del proyecto se ajusta a una distribución normal con la media, igual a la suma de las medias, a lo largo de la ruta con el tiempo de terminación esperado más largo y la varianza, igual a la suma de las varianzas, a lo largo de esta misma ruta. Según la configuración de la red, suponer que dos o más rutas son independientes podría dar resultados más exactos. Aunque el enfoque PERT es solamente aproximado, en realidad proporciona una medida del efecto de la incertidumbre de los tiempos de actividad sobre el tiempo total de terminación del proyecto.

También se contemplaron las *consideraciones de los recursos*. Los métodos tradicionales CPM y PERT no toman en cuenta el hecho de que los programas pueden ser no factibles debido a recursos insuficientes. Ejemplos típicos de recursos escasos que pueden dar lugar a un programa de proyecto no factible son la fuerza de trabajo, las materias primas y el equipo. Cuando los programas son no factibles, las actividades no críticas deben reprogramarse, de ser posible, dentro de la holgura disponible. Si no lo son, tal vez tendrían que retrasarse las actividades críticas y la fecha de terminación del proyecto se aplazaría. Los perfiles de carga de recursos son una herramienta útil para determinar los requerimientos puestos sobre los recursos por cualquier programa. A medida que evolucionan los proyectos grandes, dichos perfiles pueden usarse como guías preliminares de planeación.

También discutimos el *diseño organizacional para una administración efectiva de proyectos*. Desde la perspectiva de la administración de proyectos, la tradicional organización en línea es la estructura organizacional más débil. Al otro lado del espectro está la organización de proyectos. Aquí, se libera a los empleados de sus responsabilidades habituales durante la vida del proyecto. Algunas compañías han experimentado con las organizaciones matriciales, que es un híbrido de los dos diseños. Sin embargo la mayoría de las compañías han mantenido la tradicional estructura en línea.

El capítulo concluyó con un breve panorama del software disponible para administrar proyectos. La explosión de la computadora personal ha sido paralela a una explosión de software. La administración de proyectos no es la excepción. Hay un rango completo de productos de software disponibles. La mayoría de los programas disponibles están diseñados para correr en una PC para un solo usuario. Aunque gran parte del software basado en PC puede manejar proyectos y recursos múltiples, los sistemas muy grandes requieren herramientas más poderosas. Existen programas disponibles para las computadoras *mainframe* y sistemas cliente-servidor que permiten usuarios, proyectos y recursos múltiples. Estos paquetes pueden formar una interfase con bases de datos grandes y con otras partes de los sistemas de producción y financieros de la compañía.

Problemas adicionales de programación de proyectos

26. Dos hermanos han comprado un pequeño terreno en el centro del pueblo, donde piensan construir una estación de gasolina. La estación tendrá dos bombas, un área de servicio para agua y mantenimiento de las llantas, y un edificio principal con salas de descanso, oficinas y área de cajas. Antes de que comience la excavación del sitio, las autoridades locales deben aprobar que en el sitio se construya una estación de gasolina y tener la certeza de que al colocar los tanques de almacenamiento no se va a interferir con las líneas de agua, gas y electricidad que ya están colocadas.

    Una vez que el sitio haya sido aprobado, la excavación puede comenzar. Después de la excavación, pueden comenzar las tres partes primarias de la construcción: instalación de los tanques de gasolina, la construcción del área de servicio de agua y llantas (incluyendo la instalación del compresor de aire) y la construcción del edificio principal. La conformación del terreno puede comenzar después de la terminación de todos los edificios. Después de la conformación, el sitio debe limpiarse y deben levantarse los letreros de la estación. Sin embargo, antes que la estación pueda abrir para los negocios, debe inspeccionarse, probarse y aprobarse el compresor de aire.

    Las actividades y el tiempo requerido para cada una de ellas son los siguientes:

Actividad	Tiempo requerido (semanas)
A: Obtener la aprobación del sitio	4
B: Comenzar la excavación del sitio	2
C: Colocar y asegurar los tanques de gasolina	3
D: Instalar las bombas de gasolina	1
E: Conectar y probar las bombas de gasolina	1
F: Construir el área de servicio	2
G: Instalar y conectar el agua y el compresor de aire	3
H: Probar el compresor	1
I: Construir el edificio principal incluyendo las salas de descanso, la oficina y el área de cajas	5
J: Instalar la plomería y las conexiones eléctricas en el edificio principal	3
K: Cubrir los tanques y conformar el área	4
L: Limpiar el lugar	2
M: Levantar los letreros de la estación	1

a) A partir de la descripción del proyecto, determine las relaciones de precedencia de actividad y desarrolle una red para el proyecto.

b) Establezca la ruta crítica y los tiempos de inicio y de terminación más próximos y más lejanos para cada actividad.

c) Dibuje el diagrama de Gantt para este proyecto basándose en los primeros tiempos.

d) Suponga que el compresor de aire no funciona correctamente y debe reemplazarse. Toma dos semanas obtener otro compresor y probarlo. ¿Se retrasará necesariamente el proyecto por esta causa?

e) Enliste las actividades que deben terminarse hacia el final de la semana 15 con la finalidad de garantizar que el proyecto no se retrase.

f) Resuelva este problema con el uso de programación lineal.

27. Se está filmando una escena para una película. Se ha identificado un total de 11 actividades diferentes para la filmación. Primero, se verifica la continuidad del guión cinematográfico, se levanta y decora el escenario, y se maquilla a los actores. Después de montar el escenario, la iluminación se coloca en su lugar. Los actores se visten una vez que se les ha maquillado. Cuando se han realizado estas cinco actividades, comienza el primer ensayo de la escena, que es seguido por un segundo ensayo con el camarógrafo. Mientras que transcurren los ensayos, se hacen verificaciones del equipo de audio y de video. Después de terminar tanto los ensayos como las verificaciones, se filma la escena. Finalmente el director la ve para determinar si es necesario filmarla otra vez.

La lista de actividades y los tiempos de actividad son

Actividad	**Tiempo requerido (días)**
A: Revisar la continuidad de la historia en el guión cinematográfico	2.0
B: Decorar el escenario; colocar los puntales necesarios	4.5
C: Revisar la iluminación de la escena	1.0
D: Aplicar el maquillaje a los actores	0.5
E: Vestimentas para los actores	1.5
F: Primer ensayo (solamente actores)	2.5
G: Verificación del video	2.0
H: Verificación del sonido	2.0
I: Segundo ensayo (con cámara y luces)	2.0
J: Filmación de la escena	3.5
K: Aprobación de la escena por el director	1.5

a) Desarrolle una red para la filmación de la escena.

b) Calcule los primeros y últimos tiempos de terminación y de inicio para cada una de las actividades, e identifique la ruta crítica.

c) Dibuje el diagrama de Gantt para este proyecto. Suponga que se programan actividades con holgura de modo que haya la misma holgura antes y después de las actividades.

d) Suponga que la verificación del video (G) muestra que el equipo está defectuoso. Se requieren cuatro días adicionales para obtener y probar equipo nuevo. ¿Cuál será el atraso que resulte en el tiempo total necesario para filmar el proyecto?

e) Una de las vestimentas se daña cuando se ajusta (actividad E). ¿Cuánto tiempo extra está disponible para remiendos sin retrasar el proyecto?

f) ¿Qué tipo de retrasos puede usted imaginar como resultado de la incertidumbre en el tiempo de la actividad K?

g) Resuelva este problema con programación lineal.

28. Se construye un sistema de guía y de detección como parte de un proyecto grande de defensa. La detección consta de subsistemas de radar y de sonar. Se requiere equipo separado para cada uno de los subsistemas. En cada caso, el equipo debe calibrarse antes de la producción. Después de la producción, cada subsistema se prueba en forma independiente. El radar y el sonar se combinan para formar el sistema de detección, que también debe probarse antes de la integración al sistema de guía. La prueba final del sistema completo requiere de un equipo complejo. Las actividades y los tiempos de actividad son

Actividad	**Tiempo requerido (días)**
A: Calibrar la máquina 1 (para el radar)	2.0
B: Calibrar la máquina 2 (para el sonar)	3.5
C: Calibrar la máquina 3 (para la guía)	1.5
D: Ensamblar y preparar el equipo para la prueba final	7.0
E: Hacer el subsistema de radar	4.5
F: Hacer el subsistema de sonar	5.0
G: Hacer el subsistema de guía	4.5
H: Probar el subsistema de radar	2.0
I: Probar el subsistema de sonar	3.0
J: Probar el subsistema de guía	2.0
K: Ensamblar el subsistema de detección (radar y sonar)	1.5
L: Probar el subsistema de detección	2.5
M: Ensamblado final de los tres sistemas	2.5
N: Prueba del ensamblado final	3.5

a) Construya una red para este proyecto.

b) Determine los tiempos de inicio y terminación más próximos y más lejanos para todas las actividades e identifique la ruta crítica.

c) Dibuje un diagrama de Gantt para este proyecto con base en los tiempos más próximos.

d) ¿De cuánto tiempo se dispone para ensamblar y calibrar el equipo de prueba final sin retrasar el proyecto?

e) ¿Cuáles son las actividades que deben terminarse hacia el final de 10 días para garantizar que el proyecto no se retrase?

f) ¿Cuáles son las actividades que deben iniciarse hacia el final de 10 días para garantizar que el proyecto no se retrase?

g) Revuelva este problema usando programación lineal.

29. Considere la filmación de la escena descrita en el problema 27. Basándose en la experiencia pasada, el director no confía mucho en las estimaciones de tiempo de algunas de las actividades. Las estimaciones del director para los tiempos mínimo, más probable y máximo para estas actividades son

Actividad	***a***	***m***	***b***
F	2	3	12
I	1	2	8
J	3	4	10
K	1	2	7

a) Incluyendo estas estimaciones de tiempo PERT, ¿cuánto tiempo se espera que dure la filmación?

b) ¿Cuál es la probabilidad de que el número de días requeridos para terminar la filmación de la escena sea cuando menos 30 por ciento mayor que la respuesta que usted encontró en el inciso *a*?

c) ¿Para cuántos días debe planear el director con objeto de tener un 95 por ciento de confianza de que la filmación de la escena se termine?

30. Considere los siguientes datos del tiempo y costo del proyecto:

Actividad	Predecesores inmediatos	Tiempo normal	Tiempo acelerado	Costo normal	Costo agilizado
A	—	6	6	$ 200	$ 200
B	A	10	4	600	1 000
C	A	12	9	625	1 000
D	B	6	5	700	800
E	B	9	7	200	500
F	C, D	9	5	400	840
G	E	14	10	1 000	1 440
H	E, F	10	8	1 100	1 460

a) Desarrolle una red para este proyecto.

b) Calcule los tiempos de inicio y terminación más próximos y más lejanos para cada una de las actividades. Encuentre el tiempo de holgura para cada actividad e identifique la ruta crítica.

c) Suponga que los costos indirectos del proyecto ascienden a $200 por día. Encuentre el número óptimo de días para ejecutar el proyecto acelerando un día a la vez. ¿Cuál es el costo total del proyecto para la solución óptima? ¿Qué ahorro se ha obtenido al agilizar el proyecto?

d) Resuelva este problema usando programación lineal.

Apéndice 9-A

Glosario de términos para el capítulo 9

a = Tiempo mínimo de actividad para PERT.

b = Tiempo máximo de actividad para PERT.

EF_i = Tiempo de terminación más próximo de la actividad i.

ES_i = Tiempo de inicio más próximo de la actividad i.

LF_i = Tiempo de terminación más lejano de la actividad i.

LS_i = Tiempo de inicio más lejano de la actividad i.

m = Tiempo más probable de actividad para PERT.

M_{ij} = Tiempo acelerado de la actividad (ij).

μ = Estimación del tiempo esperado de la actividad para PERT.

N_{ij} = Tiempo normal de la actividad (ij).

σ = Estimación de la desviación estándar del tiempo de actividad para PERT.

T = Tiempo de terminación del proyecto para PERT; T es una variable aleatoria.

t_{ij} = Tiempo requerido para la actividad (ij); t_{ij} es una constante en la formulación estándar y una variable en la formulación costo-tiempo.

x_i = Tiempo de inicio más próximo para el nodo i (formulación de programación lineal).

Bibliografía

Adamiecki, Karol. "Harmonygraph". *Polish Journal of Organizational Review*, 1931. (En polaco.)

Archibald, R. D. y R. L. Villoria. *Network-Based Management Systems (PERT/CPM).* Nueva York: John Wiley & Sons, 1967.

Burgess, A. R. y J. B. Killebrew. "Variational Activity Level on a Cyclic Arrow Diagram". *Journal of Industrial Engineering* 13, 1962, pp. 76-83.

Charnes, A. y W. W. Cooper. "A Network Interpretation and a Directed Subdual Algorithm for Critical Path Scheduling". *Journal of Industrial Engineering* 13, 1962, pp. 213-219.

Cooprider, D. H. "Overview of Implementing a Project Control System in the Nuclear Utility Industry". *Cost Engineering* 36, núm. 3, marzo de 1994, pp. 21-24.

Davis, E. W. y J. H. Patterson. "A Comparison of Heuristic and Optimum Solutions in Resource-Constrained Project Scheduling". *Management Science* 21, 1975, pp. 944-955.

Department of the Navy, Special Projects Office, Bureau of Ordnance. "PERT: Program Evaluation Research Task". Phase I Summary Report. Washington, D.C., julio de 1958.

Goodman, L. J. *Project Planning and Management.* Nueva York: Van Nostrand Reinhold, 1988.

Grubbs, F. E. "Attempts to Validate Certain PERT Statistics, or 'Picking on Pert'". *Operations Research* 10, 1962, pp. 912-915.

Kulkarni, V. G. y V. G. Adlakha. "Markov and Markov-Regenerative PERT Networks". *Naval Research Logistics Quarterly* 34, 1986, pp. 769-781.

Lockyer, K. G. *Critical Path Analysis, Problems and Solutions.* Londres: Isaac Pitman and Sons, Ltd., 1966.

Moder, J. J., C. R. Phillips y E. W. Davis. *Project Management with CPM*, *PERT, and Precedence Diagramming.* 3a. ed. Nueva York: Van Nostrand Reinhold, 1983.

Ouelette, T. "Project Management Helps Airline Stick to Schedule". *Computerworld* 28, núm. 45, 7 de noviembre de 1994, p. 79.

Patterson, J. H. "A Comparison of Exact Approaches for Solving the Multiple Constrained Resource Project Scheduling Problem". *Management Science* 30, 1984, pp. 854-867.

Sasieni, M. "A Note on PERT Times". *Management Science* 32, 1986, pp. 1652-1653.

Sullivan, R. *S.,* J. C. Hayya y R. Schaul. "Efficiency of the Antithetic Variate Method for Simulating Stochastic Networks". *Management Science* 28, 1982, pp. 563-572.

Vollmann, T. E., W. L. Berry y D. C. Whybark. *Manufacturing Planning and Control Systems.* 3a. ed. Homewood, IL: Richard D. Irwin, 1992.

Walker, M. R. y J. S. Sayer. "Project Planning and Scheduling". Reporte 6959. Wilmington, DE: E. I. du Pont de Nemours & Co., Inc., marzo de 1959.

Wiest, J. D. y F. K. Levy. A *Management Guide to PERT/CPM,* 2a. ed. Englewood Cliffs, NJ: Prentice Hall, 1977.

Capítulo **Diez**

Distribución y ubicación de instalaciones

Panorama general del capítulo

Objetivo

Entender los principales aspectos que enfrenta una compañía al diseñar y ubicar instalaciones nuevas, y aprender las técnicas cuantitativas para ayudar en la toma de decisiones.

Puntos clave

1. *Fundamentos.* Antes de elegir la distribución apropiada de una nueva instalación, sea una fábrica, hospital, parque temático, o cualquier otra cosa, deben estudiarse primero los patrones de flujo. El más sencillo es el flujo en línea recta, como de una línea de ensamblado. Otros incluyen el flujo en U, en L, en serpentina, circular, y en S. Otro aspecto es la conveniencia o inconveniencia de ubicar las operaciones cercanas entre sí. Por ejemplo, en un hospital, la sala de emergencia debe estar cerca de la entrada al hospital, y la sala de maternidad debe estar cerca del área de cuidados para bebés prematuros. Una técnica gráfica para representar la conveniencia relativa de ubicar dos instalaciones cercanas entre sí es la de relación de actividad. Las tablas de-hacia proporcionan las distancias entre las actividades, utilizables para calcular los costos asociados con diferentes distribuciones.
2. *Tipos de distribuciones.* En la disposición de la fábrica, el tipo apropiado de distribución dependerá del entorno de la manufactura y de las características del producto. La *distribución de posición fija* es apropiada cuando se construyen artículos grandes tales como aviones o barcos, difíciles y costosos de mover. Las estaciones de trabajo se ubican alrededor del objeto, que permanece estacionario. Es más común la *distribución de producto*, donde las máquinas o estaciones de trabajo se organizan según la secuencia de operaciones requerida para elaborar el producto. Las *distribuciones de producto* son muy comunes para la producción en masa. En el caso de compañías de tamaño pequeño a medio, una distribución de proceso tiene más sentido. Se agrupan máquinas similares o procesos similares. Finalmente, las *distribuciones basadas en tecnología de grupo* podrían ser adecuadas. En este caso, las máquinas pueden agruparse en células donde cada una corresponde a una familia de partes o a un grupo de familias de partes.
3. *Técnicas computarizadas de distribución.* Para fábricas complejas grandes o instalaciones de servicios, no es práctico determinar manualmente la mejor distribución. Existen varias técnicas computarizadas de distribución que pueden ayudar a tal función. Éstas incluyen CRAFT, COFAD, ALDEP, CORELAP y PLANET. Todos estos métodos están diseñados para la disposición de la fábrica y comparten el objetivo de minimizar los costos de manejo de materiales. Tanto CRAFT como COFAD se basan en el principio de mejoramiento. Esto

implica que el usuario debe especificar una distribución inicial. A partir de ahí, se consideran intercambios por pares de los departamentos y se escoge aquel que tenga la mejora más grande.

ALDEP, CORELAP y PLANET son rutinas de construcción antes que de mejoramiento. Las distribuciones se determinan a partir del nivel de planeación inicial, y no es necesario que el usuario especifique una distribución inicial. Existe alguna controversia con respecto a si los planificadores humanos o los programas de computadora producen mejores distribuciones. En un estudio donde grupos de 20 escogidos de entre 74 personas entrenadas en técnicas de distribución se compararon con las distribuciones computarizadas, a los humanos les fue mucho mejor. Otros criticaron este estudio argumentando que la mayoría de los departamentos encargados de la distribución no tienen tan buen personal.

4. *Sistemas flexibles de manufactura.* Un sistema flexible de manufactura (SFM o FMS, *flexible manufacturing system*) es un grupo de máquinas controladas numéricamente, conectadas por un sistema de flujo de materiales controlado por computadora. Los más comunes son los usados para corte de metales y operaciones de formado así como ciertas operaciones de ensamble. Como las máquinas pueden programarse, puede usarse el mismo sistema para producir una variedad de partes diferentes. Tienden a ser muy caros (algunos cuestan más de 10 millones de dólares). Como resultado, tal vez la flexibilidad añadida no justifique el costo. Aun cuando el FMS puede tener muchas ventajas (reducción del inventario del trabajo en proceso, uso creciente de las máquinas, flexibilidad), estas ventajas se justifican rara vez, por el alto costo de estos sistemas. Una alternativa que es más popular son las células flexibles de manufactura. Estos son sistemas más pequeños que los totalmente desarrollados, pero todavía suministran mayor flexibilidad que el equipo de una sola función.
5. *Ubicación de nuevas instalaciones.* La ubicación de una nueva instalación es un problema complejo y estratégicamente importante. Los hospitales necesitan estar cerca de los centros de población de alta densidad, y los aeropuertos necesitan estar cerca de las ciudades grandes, aunque no demasiado cerca por la contaminación acústica. Frecuentemente las fábricas nuevas se localizan fuera de Estados Unidos para aprovechar los costos más bajos de mano de obra en el extranjero. Pero estos ahorros pueden tener un alto precio. La inestabilidad política, las tasas de cambio poco favorables, las deficiencias en la infraestructura, y los prolongados tiempos de demora son algunos de los problemas que surgen al ubicar las instalaciones en el extranjero. Frecuentemente estas decisiones son más estratégicas que tácticas y requieren la ponderación cuidadosa de las ventajas y desventajas al nivel de gerencia general.

   Sin embargo, en los casos donde el objetivo primario es ubicar una instalación para que esté lo más cercana posible a su base clientelar, los métodos cuantitativos pueden ser muy útiles. En estos casos, debe especificarse cómo se mide la distancia. La distancia en línea recta (también conocida como distancia euclidiana) mide la distancia más corta entre dos puntos. Sin embargo, la distancia en línea recta no siempre es la medida más apropiada. Por ejemplo, cuando se ubica una estación de bomberos, debe considerarse la distribución de las calles. El uso de la distancia rectilínea (tal como se mide solamente con movimientos horizontales y verticales) tendría más sentido en este contexto. Otra consideración es que no todos los clientes son del mismo tamaño. Por ejemplo, una panadería haría entregas más grandes a un supermercado o a un almacén que a una tienda de abarrotes. Aquí se usaría un criterio de distancia ponderada. En lo que resta de este capítulo, revisamos varias técnicas cuantitativas para encontrar la mejor ubicación de una instalación individual bajo varios objetivos.

La ubicación de instalaciones y el diseño eficiente de esas instalaciones son aspectos estratégicos importantes para los negocios así como para las instituciones militares, las no lucrativas, y el gobierno. Durante la década de 1980, Silicon Valley, en el norte de California, experimentó un gran crecimiento en la microelectrónica e industrias relacionadas. Puede tenerse alguna idea

de qué tan importante fue este crecimiento al consultar el Directorio de Miembros de la American Electronics Association: más del 50 por ciento de los que están en la lista se encuentran en el área de la bahía. La mayoría de estas compañías fueron "negocios nuevos", que adaptaron una nueva tecnología a un segmento especializado del mercado. Con el surgimiento de la demanda de microcomputadoras, aparecieron industrias de apoyo que producen unidades de disco duro, discos flexibles, equipo de fabricación de semiconductores, redes de área locales, y un alud de productos relacionados. En muchos casos, se requirieron grandes inversiones en equipo de capital antes de poder vender la primera unidad. Un ejemplo típico es la corporación Read-Rite de Milpitas, California, el mayor productor independiente en el mundo de cabezas para película delgada de lectura y escritura destinadas a las unidades de disco duro. Cuando se fundó la compañía en 1983, se requirió una inversión inicial de capital de $40 millones. Los inversionistas asumieron la posibilidad de que la compañía no pudiera sobrevivir. Read-Rite es un caso típico entre cientos de negocios nuevos de alta tecnología.

Algunos aspectos fundamentales de la estrategia que enfrenta cualquier organización manufacturera consisten en decidir qué equipo debe comprarse, cómo deben organizarse las instalaciones y dónde deben ubicarse las mismas. Las industrias de servicios también encaran el problema de encontrar distribuciones efectivas. Alcanzar flujos eficientes de pacientes en los hospitales es un ejemplo. Otro ejemplo es el cuidado obvio que se usa al diseñar la distribución de muchos parques temáticos a lo largo de Estados Unidos. Quienquiera que haya visitado el Zoológico de San Diego o Disney World en Orlando, Florida, puede apreciar la importancia de una distribución efectiva y de un manejo eficiente de las personas en las industrias de servicios.

Tompkins y White (1984) estimaron que se ha gastado un 8 por ciento del producto nacional bruto de Estados Unidos en instalaciones nuevas anualmente desde 1955. Esto no incluye el costo de modificación de las instalaciones existentes. Si las estimaciones de los autores son correctas, los estadounidenses están gastando más de 500 mil millones de dólares anualmente en construir y modificar instalaciones. Sostienen que de 20 a 50 por ciento de los gastos totales de operación en manufacturas se atribuyen a los costos del manejo de materiales. Asimismo, afirman que la planeación efectiva de instalaciones podría reducir estos costos entre 10 y 30 por ciento anualmente.

Una parte importante del enorme éxito de las compañías japonesas en conseguir el dominio manufacturero en varias áreas clave es la producción eficiente. Ésta incluye el diseño eficiente del producto, la participación de los empleados, los sistemas livianos de manejo de materiales de inventario, y la distribución y organización inteligentes de las instalaciones.

En el capítulo 1 se trataron algunos de los aspectos cualitativos que la gerencia debe considerar al decidir la ubicación de las nuevas instalaciones. Este capítulo trata con mayor profundidad los aspectos asociados con la ubicación de nuevas instalaciones y los métodos para determinar la mejor distribución de operaciones en dichas instalaciones. Además de explorar los factores cualitativos que deben tomarse en cuenta, también consideraremos cómo un gerente puede emplear las computadoras y las técnicas cuantitativas para ayudar a estas decisiones complejas.

La forma en que debe diseñarse la distribución de un centro de trabajo es en cierto sentido una versión especial del problema de localización. La determinación de una distribución adecuada implica encontrar las ubicaciones de los departamentos dentro de algún límite específico. En el diseño de nuevas instalaciones, el planeador también debe decidir el tamaño y la forma de la instalación así como la configuración de los departamentos que están dentro.

Las técnicas cuantitativas son muy útiles cuando la meta es minimizar o maximizar un objetivo dimensional individual como un costo o una ganancia. La función objetivo que se usa en los problemas de ubicación generalmente incluye la distancia euclidiana o bien la rectilínea (estos términos se definirán posteriormente en la sección 10.8). Sin embargo, la reducción de la distancia total recorrida puede no tener sentido en muchos casos. Como un ejemplo extremo, considere el problema de ubicar una escuela. Es evidente que una ubicación que requiere que 100 estudiantes viajen 16 kilómetros cada uno no es igualmente conveniente que una que requiere que 99 estudiantes viajen kilómetro y medio, y uno viaje 145.

Para los problemas de distribución, el objetivo más común que se usa en los modelos matemáticos es minimizar el costo del manejo de los materiales. Además, estos modelos invariablemente suponen que se conoce con certeza el número de viajes en el manejo de materiales desde cada uno de los centros de trabajo a los demás. En la mayoría de los escenarios reales estas hipótesis, en el mejor de los casos, son ingenuas. Otra deficiencia de los modelos matemáticos es que ignoran factores tales como la seguridad de la planta, la flexibilidad de la distribución para cambios futuros del diseño, el ruido y la estética.

Esto no implica que los modelos matemáticos y computacionales no sean útiles para resolver los problemas de distribución y de localización. Lo que significa es que las soluciones cuantitativas no deben aceptarse a ciegas. Deben considerarse cuidadosamente en el contexto del problema. Si se usan de modo apropiado, los resultados de los modelos matemáticos y computacionales pueden reducir significativamente el número de alternativas que el analista debe considerar.

Aplicación instantánea

SUN MICROSYSTEMS ES EL PIONERO DEL NUEVO FLEX OFFICE SYSTEM

Sun Microsystems, con base en Santa Clara, California, se fundó en 1982, y en dos décadas se ha transformado en un proveedor a nivel mundial de productos de hardware y software con ventas que sobrepasan los $10 millardos al año. Un problema encarado por Sun, y otras compañías con base en Silicon Valley, es el alto costo de los bienes raíces. Por ejemplo, actualmente los costos habitacionales son más elevados en varias partes del área de la bahía que en cualquier otra parte de Estados Unidos, incluyendo los mercados históricamente más caros de Nueva York, Boston y Honolulu. Aun cuando la quiebra del mercado de valores de alta tecnología de 2000 ha amortiguado un poco el mercado comercial de los bienes raíces, estos costos todavía son bastante altos si los comparamos con prácticamente cualquier otra área metropolitana.

Para ayudar a combatir estos costos, Sun se ha comprometido con un programa innovador de asignaciones de oficina flexibles para sus empleados. El modelo tradicional de negocios, que siguen virtualmente todas las compañías del mundo, es que a cada empleado se le asigne una oficina personal. ¿Cuál es el uso de estas oficinas? Esto depende, por supuesto, de la función industrial y laboral de la persona. De acuerdo con Crawford Beveridge, el jefe de recursos humanos de Sun, del 30 al 35 por ciento de la mayoría de los empleados de las compañías no está en sus oficinas a cualquier hora del día. ¿Por qué? Las razones varían, pero se incluyen clases, conferencias, visitas a los clientes, incapacidades, vacaciones, juntas extramuros, y el teletrabajo (el caso del empleado que hace su trabajo en casa en una computadora personal o una terminal que comunica con la oficina empleadora). Lo que significa que aproximadamente un tercio de los bienes raíces de la mayoría de las compañías no se utiliza. En el caso de Sun, es probable que esta proporción sea todavía mayor, puesto que se permite a muchos empleados que trabajen en su casa para evitar los largos lapsos del viaje de ida y vuelta.

Para mejorar la utilización del espacio Sun desarrolló un programa innovador. La mayoría de los empleados no tienen una oficina única asignada. Los empleados que son parte del así llamado sistema iWork pueden reservar con antelación una oficina por un tiempo específico en alguna de varias localidades. Tendrían que llevar con ellos los efectos personales necesarios. Lo que realmente hace funcionar a este sistema es una tarjeta que los empleados pueden enchufar en cualquier estación de trabajo de Sun para tener acceso inmediato a todos sus archivos. Las tarjetas de iWork suministran una medida de seguridad extra ya que los empleados elegibles deben tener tanto el hardware (las tarjetas) como el software (las claves de acceso) correctos para tener acceso a sus archivos. De acuerdo con Mark Tolliver, funcionario de estrategia principal de Sun. "Con las oficinas flexibles en San Francisco, por ejemplo, en lugar de hacer rondas de ida y vuelta en la autopista 101 para llegar al trabajo, nuestros empleados con base en San Francisco pueden caminar un poco o usar el transporte público hasta la oficina de San Francisco. Mediante el uso de nuestra tarjeta de acceso, ellos pueden ser completamente productivos con un método hecho a la medida. Es un gran cambio, pero yo pienso que representa una línea de meta hacia la cual veremos que muchas organizaciones se mueven."[1]

La meta de Sun en el programa iWork es tener 1.8 empleados para cada oficina y reducir dramáticamente sus costos generales en bienes raíces. ¿Qué tan bien está trabajando el programa? Primero, Beveridge estima que Sun ha podido retener a 680 empleados que habría perdido al haber otorgado a estos empleados la opción del teletrabajo. Segundo, también estima que Sun está ahorrando aproximadamente $50 millones al año en costos de bienes raíces y prevé que cuando iWork esté totalmente implementado va a ahorrar a la compañía $140 millones al año.

[1] Schauer, J., "Sun Rises on Business Intelligence", *DM Review* 13, núm. 10 (octubre de 2003).

10.1 EL PROBLEMA DE LA DISTRIBUCIÓN DE INSTALACIONES

Determinar la mejor distribución de una instalación es un problema clásico en la ingeniería industrial. Con frecuencia a los primeros ingenieros industriales se les conocía como expertos en eficiencia y estaban interesados en determinar distribuciones para optimizar alguna medida de la eficiencia de la producción. En la mayoría de los casos, este punto de vista sigue vigente hoy, especialmente en el diseño de plantas. Sin embargo, los problemas de distribución se presentan en muchos ambientes fuera de la planta. Algunos de éstos incluyen

1. Hospitales
2. Almacenes
3. Escuelas
4. Oficinas
5. Estaciones de trabajo
6. Bancos
7. Centros comerciales
8. Aeropuertos
9. Plantas industriales

Cada uno de estos problemas de distribución tiene características únicas. En este capítulo nos centraremos en las técnicas para encontrar las distribuciones de una planta industrial, aunque muchos de los métodos también pueden aplicarse a otras instalaciones. Los objetivos de un estudio de distribución de planta podrían incluir uno o más de los siguientes:

1. Minimizar la inversión en equipo nuevo.
2. Minimizar el tiempo en la producción.
3. Utilizar el espacio con la mayor eficiencia.
4. Tener en mente la conveniencia, la seguridad y la comodidad de los empleados.
5. Conservar una disposición flexible.
6. Minimizar el costo de manejo de materiales.
7. Facilitar el proceso de manufactura.
8. Facilitar la estructura organizacional.

Vollmann y Buffa (1966) sugieren nueve pasos como una guía de análisis de los problemas de distribución:

1. Determine la compatibilidad de los modelos de distribución de manejo de materiales con el problema bajo estudio. Encuentre todos los factores que puedan modelarse como flujo de materiales.
2. Determine las subunidades básicas para el análisis. Determine la definición apropiada de un departamento o de una subunidad.
3. Si debe usarse un modelo matemático o computacional, determine la compatibilidad de la naturaleza de los costos en el problema y en el modelo. Es decir, si el modelo supone que los costos de manejo de los materiales son lineales e incrementales (como ocurre en la mayoría), determine si estos supuestos son realistas.
4. ¿Qué tan sensible es la solución a los supuestos de los datos de flujos? ¿Cuál es el impacto de los cambios aleatorios en estos datos?
5. Reconozca las idiosincrasias del modelo y trate de encontrar mejoras.
6. Examine los aspectos a largo plazo asociados con el problema y las implicaciones a largo plazo de la solución propuesta.
7. Considere el problema de distribución como un problema de sistemas.

8. Pondere la importancia de los factores cualitativos.
9. Seleccione las herramientas apropiadas para el análisis.

10.2 PATRONES DE FLUJO

Como observamos en la sección 10.1, el objetivo más frecuentemente empleado en el análisis cuantitativo de los problemas de distribución es minimizar el costo de manejo de materiales. Cuando éste es el objetivo primario, se necesita un análisis de flujo de la instalación. Los patrones de flujo pueden clasificarse como horizontales o verticales. Un patrón de flujo horizontal resulta apropiado cuando todas las operaciones se ubican en el mismo piso, y un patrón vertical es el adecuado cuando las operaciones están en estructuras de niveles múltiples. Francis y White (1974) mencionan seis patrones de flujo horizontal y seis patrones de flujo vertical. Los seis patrones horizontales aparecen en la figura 10-1.

El patrón más simple es (*a*), que es un flujo en línea recta. La principal desventaja de este patrón consiste en que se requieren andenes separados y personal para recibir y embarcar bienes. La forma en L se usa para remplazar el flujo en línea recta si la configuración del edificio o la línea así lo necesitan. La forma en U aventaja a la configuración en línea recta en que permite que el embarque y la recepción tengan la misma ubicación. El patrón circular es similar a la forma en U. Los dos patrones restantes se emplean cuando el espacio requerido para las operaciones de producción es demasiado grande como para usar otros patrones.

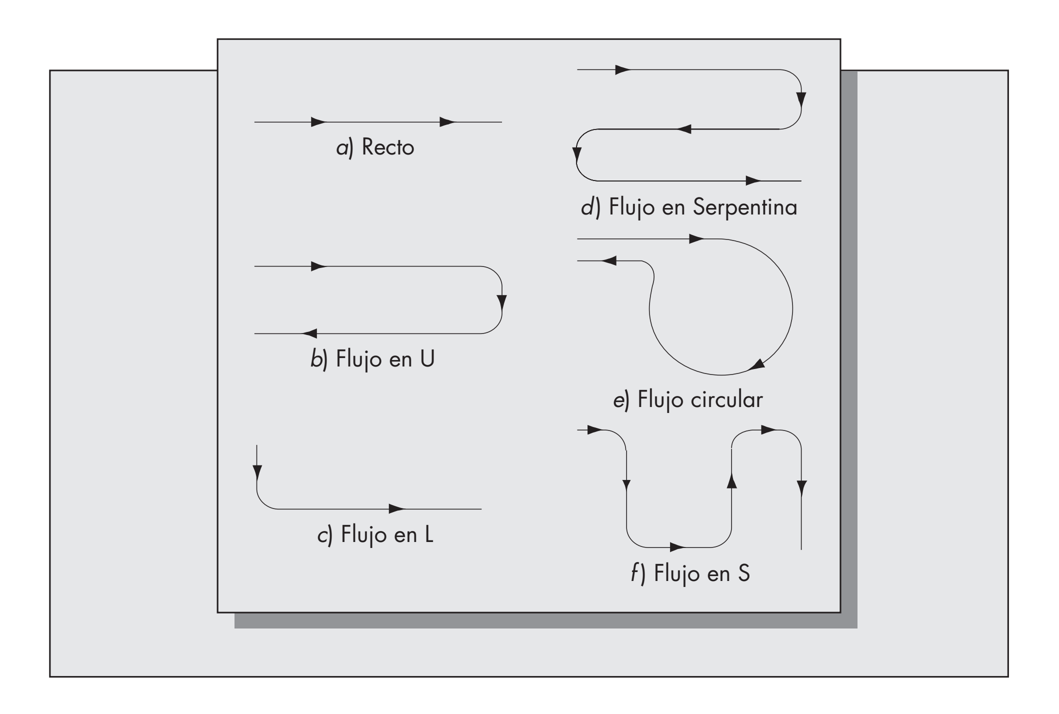
Dos diagramas que suministran información útil con respecto a los flujos son 1) el diagrama de relación de actividades, y 2) el diagrama de-hacia.

Diagrama de relación de actividades

Un diagrama de relación de actividades (también llamado diagrama de relación para abreviar) es un medio gráfico con el que se representa la conveniencia de ubicar pares de operaciones cercanas entre sí. Se han sugerido los siguientes códigos de letras para determinar una clasificación de "cercanía".

A Absolutamente necesario. Como dos operaciones pueden usar el mismo equipo o las mismas instalaciones, deben localizarse cercanas entre sí.

FIGURA 10-1
Seis patrones de flujo horizontal.



E Especialmente importante. Las instalaciones pueden requerir el mismo personal o los mismos registros, por ejemplo.

I Importante. Las actividades pueden ordenarse secuencialmente en el flujo de trabajo normal.

O Importancia ordinaria. Sería conveniente que las instalaciones estén cercanas entre sí, pero no es imprescindible.

U No es importante. No importa si las instalaciones se ubican próximas entre sí o no.

X No es deseable. La ubicación de un departamento de soldadura cerca de uno que emplea líquidos inflamables entraría en esta categoría.

Las clasificaciones de cercanía se representan en un diagrama de relación de actividades que especifica la clasificación apropiada para cada par de departamentos. Considere el siguiente ejemplo.

Ejemplo *10.1*

Meat Me, Inc., es una cadena concesionaria de restaurantes de comida rápida donde se venden hamburguesas. Un nuevo restaurante se está ubicando en una comunidad suburbana en crecimiento cerca de Reston, Virginia. Cada restaurante tiene los siguientes departamentos:

1. Cocinado de hamburguesas.
2. Cocinado de papas fritas.
3. Empaque y almacenaje de hamburguesas.
4. Despachadores de refrescos.
5. Servidores de mostrador.
6. Servidor para automovilistas.

Las hamburguesas se cocinan en una parrilla grande, y las papas fritas se fríen sumergidas en aceite caliente. Por razones de seguridad la compañía requiere que estas áreas de cocinado no estén cercanas entre sí. Todas las hamburguesas se envuelven en forma individual tras el cocinado y se almacenan cerca del mostrador. El mostrador de servicio puede alojar a seis servidores, y el sitio tiene un área reservada para una ventanilla de servicio a automovilistas.

En la figura 10–2 se presenta un diagrama de relación de actividades. A cada par de actividades se le asigna una letra: A, E, I, O, U, o X. Una vez que se determina una distribución final, puede compararse la proximidad de los diferentes departamentos con las clasificaciones de cercanía en el diagrama.

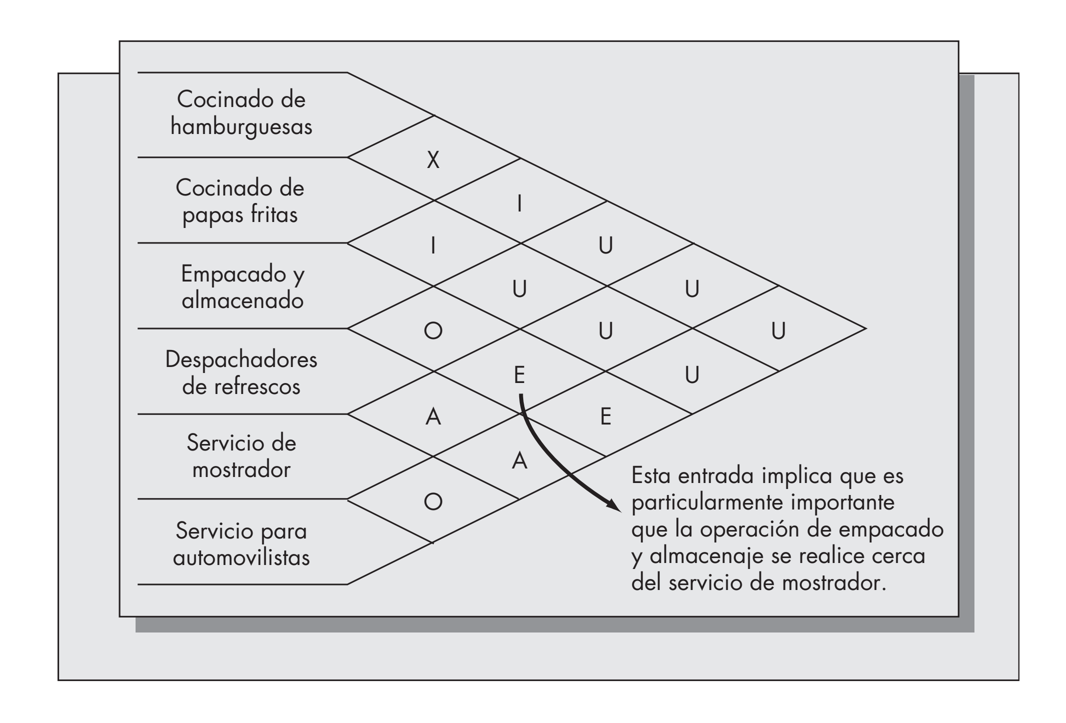


FIGURA 10-2
Diagrama de relación de actividades para el restaurante de comida rápida Meat Me.

Observe que la figura 10–2 proporciona solamente la clasificación de cercanía para cada par de departamentos. En la concepción original del diagrama, en cada celda también se incluye un número que da la razón de cada clasificación de cercanía. Dichos números no aparecen en este ejemplo.

Diagrama de-hacia

Un diagrama de-hacia es similar a la tabla de distancias que aparece en la parte inferior de muchos mapas carreteros y que indica la distancia entre pares seleccionados de ciudades. Los diagramas de-hacia se usan para analizar el flujo de materiales entre dos departamentos. Las dos formas más comunes son los diagramas que muestran las distancias entre dos departamentos y los que muestran el número de viajes de manejo de materiales por día entre dos departamentos. Un diagrama de-hacia difiere de un diagrama de relación de actividades en que el diagrama de-hacia se basa en una distribución específica. Resulta una manera conveniente de resumir los datos de flujo correspondientes a una distribución dada.

Ejemplo *10.2*

Un taller de maquinado tiene seis centros de trabajo. Los centros de trabajo contienen uno o más de los siguientes tipos de máquinas:

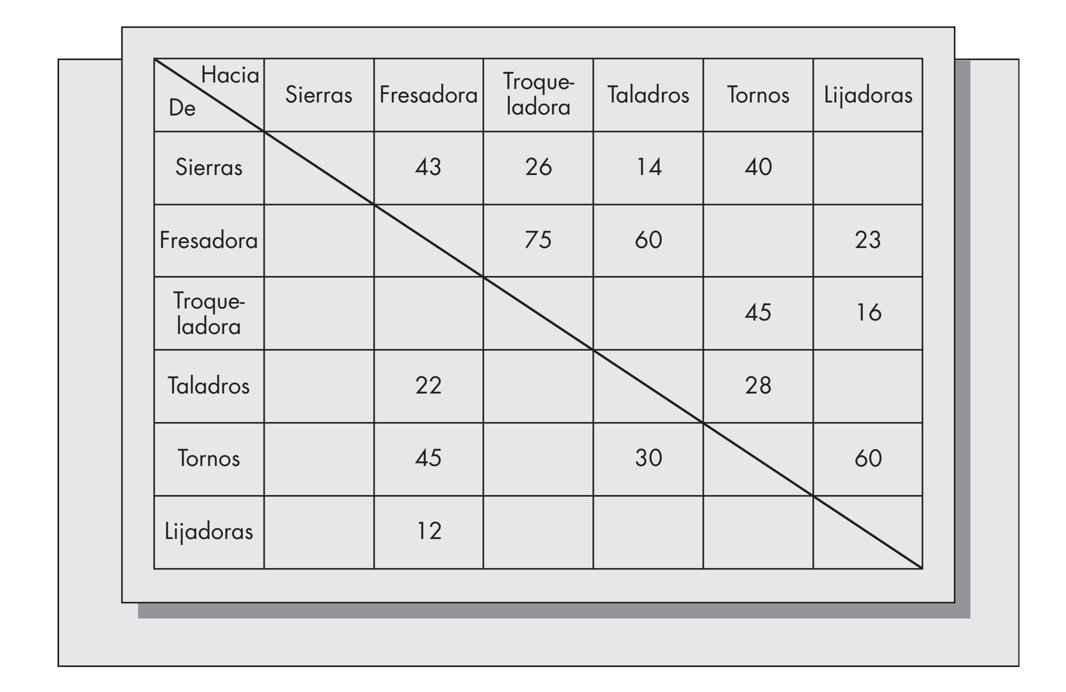
1. Sierras
2. Fresadoras
3. Troqueladoras
4. Taladros
5. Tornos
6. Lijadoras

El diagrama de-hacia de la figura 10–3 muestra la distancia en pies entre los centros de los seis departamentos. Observe que el diagrama del ejemplo es simétrico; es decir, la distancia de recorrido entre A y B es la misma que la distancia de recorrido entre B y A. Sin embargo, éste no siempre es el caso. Puede haber canales de un solo sentido, que permiten el flujo del material solamente en una dirección, o un sistema automatizado de manejo de materiales, que mueva tarimas en una sola dirección.

La figura 10–4 muestra un diagrama de-hacia que indica el número de viajes de manejo de material por día. Estas cifras podrían basarse en una mezcla específica de productos elaborados en el taller, o simplemente ser representativas de un día promedio.

FIGURA 10-3
Diagrama de-hacia que muestra las distancias entre seis centros de departamentos (medidas en pies).

Hacia / De	Sierras	Fresadora	Troque-ladora	Taladros	Tornos	Lijadoras
Sierras		18	40	30	65	24
Fresadora	18		38	75	16	30
Troque-ladora	40	38		22	38	12
Taladros	30	75	22		50	46
Tornos	65	16	38	50		60
Lijadoras	24	30	12	46	60	



De \ Hacia	Sierras	Fresadora	Troque-ladora	Taladros	Tornos	Lijadoras
Sierras		43	26	14	40	
Fresadora			75	60		23
Troque-ladora					45	16
Taladros		22			28	
Tornos		45		30		60
Lijadoras		12				

FIGURA 10-4
Diagrama de-hacia que muestra el número diario de viajes de manejo de material.

Suponga que el departamento de contabilidad en la compañía ha estimado que el costo promedio de transporte de material por pie en el taller de máquinas equivale a 20 centavos. Usando este resultado, es posible desarrollar un tercer diagrama de-hacia que exprese el costo diario promedio del manejo de materiales de cada uno de los departamentos hacia cada uno de los otros departamentos. Por ejemplo, la distancia que separa las sierras de las fresadoras en la presente distribución es de 18 pies (de la figura 10-3) y hay un promedio de 43 viajes de manejo de material entre esos dos departamentos (de la figura 10-4). Esto se traduce en un total de (43)(18) = 774 pies recorridos en un día o a un costo total de (774)(0.2) = $154.80 por día por el manejo de material entre las sierras y las fresadoras. Los costos de manejo de material para los otros pares departamentales se muestran en la figura 10-5.

FIGURA 10-5
Diagrama de-hacia que muestra el costo diario de manejo de materiales (en dólares).

De \ Hacia	Sierras	Fresadora	Troque-ladora	Taladros	Tornos	Lijadoras
Sierras		154.8	208	84	520	
Fresadora			570	900		138
Troque-ladora					342	38.4
Taladros		330			280	
Tornos		144		300		720
Lijadoras		72				

Los diagramas de-hacia no son un medio para determinar distribuciones, sino simplemente una manera conveniente de expresar las características importantes de flujo de una distribución existente. Pueden ser útiles para comparar los costos de manejo de materiales de un pequeño número de alternativas. Como los diferentes criterios del costo del manejo de materiales son relevantes, el diagrama de-hacia debe enriquecerse con información extra, como la contenida en un diagrama de relación de actividades.

10.3 TIPOS DE DISTRIBUCIONES

En distintos ambientes de manufactura se aplican diferentes filosofías de diseño de distribución. En el capítulo 1 se estudió el problema de la concordancia del ciclo de vida del producto con el ciclo de vida del proceso representado por la matriz producto-proceso (véase particularmente la figura 1–5). La esquina superior izquierda de la matriz producto-proceso corresponde a una producción de bajo volumen y poca estandarización del producto. Una estructura de producto de este tipo generalmente se caracteriza por un ambiente del tipo de taller. En un taller hay una amplia variedad de trabajos con diferentes patrones de flujo asociados con cada uno de ellos. Una imprenta comercial es un ejemplo típico de un taller con un flujo que se mueve en desorden. A medida que aumenta el volumen, el número de productos declina y los patrones de flujo se hacen más estandarizados. Una planta ensambladora de automóviles es un buen ejemplo de la manufactura de partes discretas en gran cantidad. Para un entorno de este tipo sería adecuado un enfoque diferente para el diseño de las instalaciones de producción.

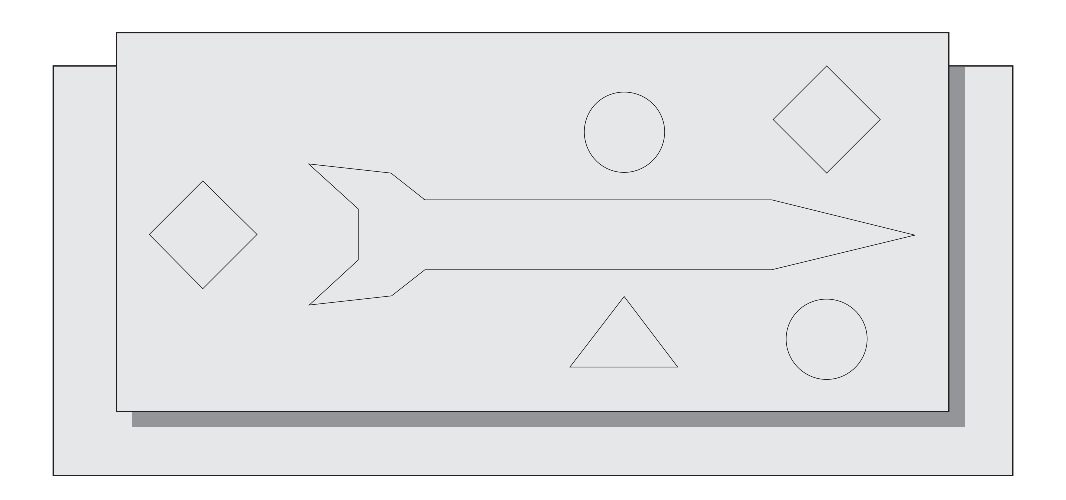
Distribuciones con posiciones fijas

Algunos productos son demasiado grandes para poder moverlos, por lo que permanecen fijos y su distribución se basa en el tamaño y la forma de éste. Ejemplos de productos que requieren una *posición fija* son los aviones grandes, los barcos y los cohetes. Para estos productos, una vez que se construye el armazón básico, las diferentes funciones requeridas se ubicarían en posiciones fijas alrededor del producto. Una distribución de proyecto tiene un concepto similar a la distribución de posición fija. Esto sería apropiado para grandes construcciones, como los edificios comerciales o los puentes. El equipo requerido se mueve al sitio y se retira cuando el proyecto se termina. En la figura 10-6 se muestra una distribución típica de posición fija.

Distribuciones por producto

En una *distribución por producto* (o distribución de flujo de producto) las máquinas están organizadas para conformar la secuencia de operaciones requerida para generar el producto. La distribución de producto es típica en la producción estandarizada de alto volumen (la esquina

FIGURA 10-6
Distribución de posición fija.



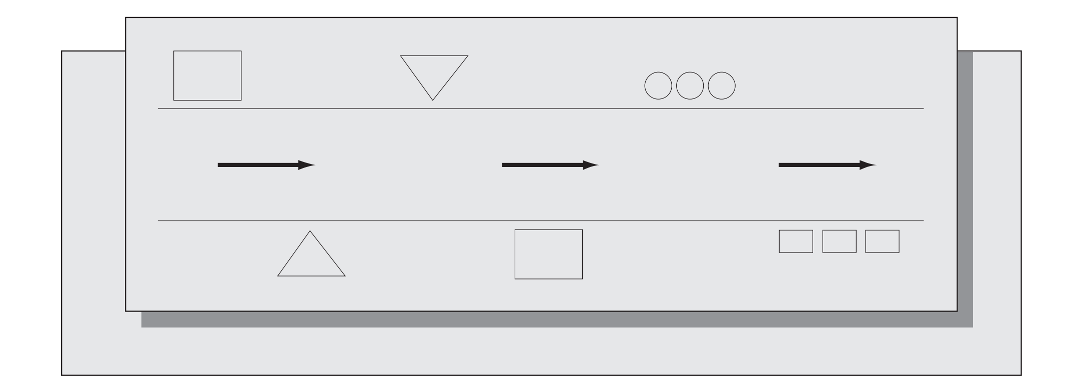


FIGURA 10-7
Distribución de producto.

inferior derecha de la figura 1-5 del capítulo 1). Una línea de ensamble (o línea de transferencia) es una distribución de producto, porque las instalaciones de ensamble se organizan según los pasos requeridos para producir el artículo. Las distribuciones de producto convienen a la producción en masa del tipo de flujo y proporcionan los tiempos de ciclo más rápidos en este entorno. Sin embargo, las líneas de transferencia son caras e inflexibles, y se vuelven estorbosas cuando se requieren cambios en el flujo del producto. Además, una línea de transferencia puede experimentar un considerable tiempo muerto. Si se detiene una parte de la línea, la línea completa podría tener que permanecer inactiva hasta que el problema se corrija. La figura 10-7 muestra una distribución de producto típica.

Distribuciones por proceso

Las distribuciones por proceso son las más comunes para fabricantes de pequeño a medio volumen. Una *distribución de proceso* agrupa máquinas similares que tienen funciones parecidas. Una distribución de proceso típica agruparía los tornos en un área, los taladros en otra área, etcétera. Las distribuciones de proceso son muy efectivas cuando hay una amplia variación en la mezcla del producto. Cada producto tiene una secuencia diferente de ruta. En un ambiente como éste sería difícil organizar a las máquinas para que se adaptaran al flujo de producción puesto que los patrones de flujo son muy variables. Las distribuciones de proceso tienen la ventaja de minimizar el tiempo muerto de máquina. Las partes provenientes de los trabajos o productos múltiples hacen cola en cada centro de trabajo para facilitar una considerable utilización de los recursos críticos. Asimismo, cuando son comunes los cambios de diseño, las rutas de las partes cambian frecuentemente. En un entorno de este tipo, la distribución de proceso permite una desorganización mínima. La figura 10–8 muestra una distribución de proceso típica. Las flechas corresponden a las rutas de las partes.

Distribuciones basadas en tecnología de grupos

Con un énfasis creciente en las fábricas automatizadas y en los sistemas flexibles de manufactura, las *distribuciones de tecnología de grupo* han recibido considerable atención en los últimos años. Para implementar una distribución por tecnología de grupos, las partes deben identificarse y agruparse con base en similitudes de la función o el diseño de la manufactura. Las partes se organizan en familias de ellas. Supuestamente, cada familia requiere un procesamiento similar, lo que sugiere una distribución basada en las necesidades de cada familia. En la mayoría de los casos, las máquinas se agrupan en células de máquinas donde cada célula corresponde a una familia específica de partes o a un grupo pequeño de familias de partes (véase la figura 10-9).

El concepto de tecnología de grupos parece ser el más adecuado para compañías grandes que producen una amplia variedad de partes para volúmenes de moderados a altos. Una compañía típica que consideraría este enfoque puede tener hasta 10000 números diferentes de par-

FIGURA 10-8
Distribución por proceso.

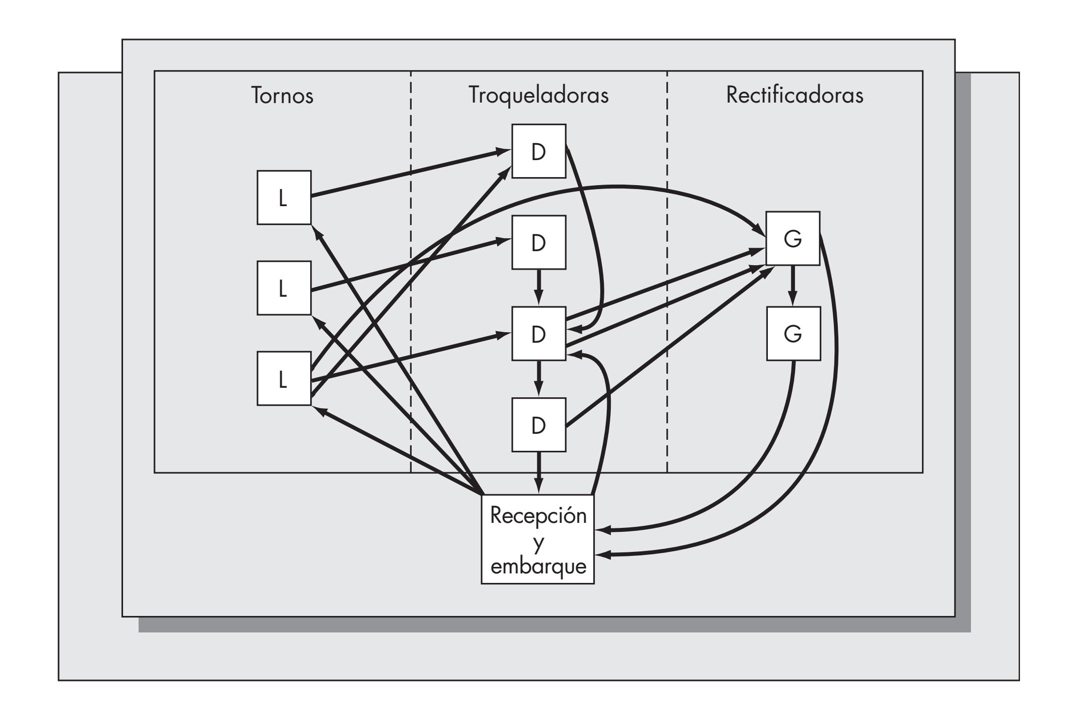
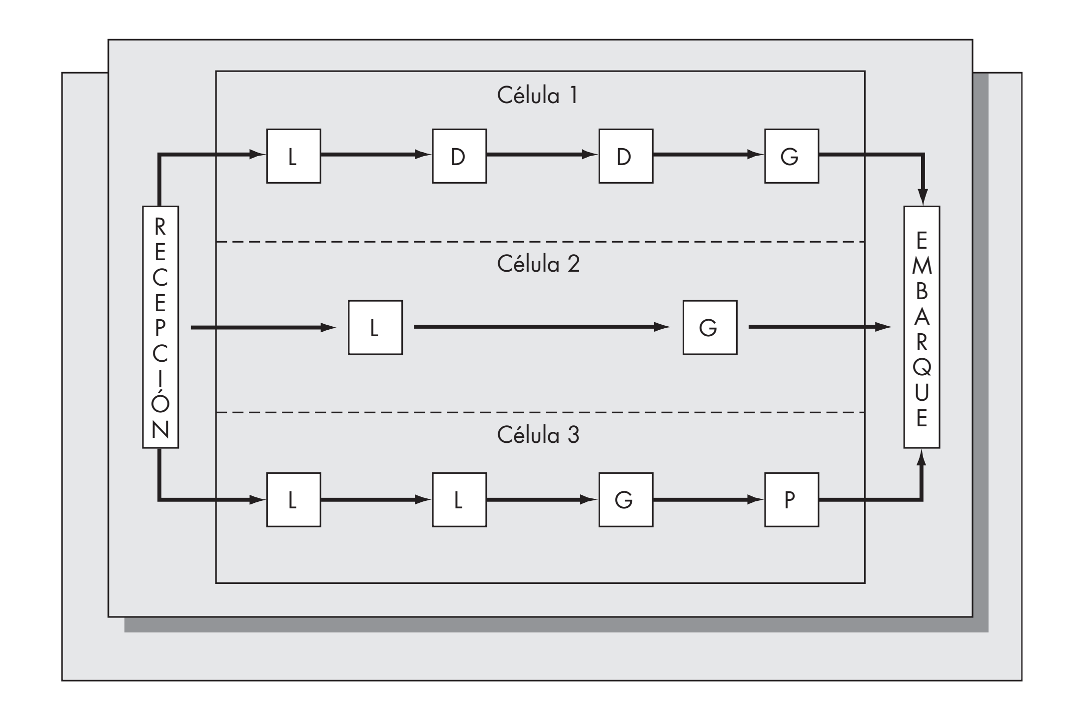


FIGURA 10-9
Distribución por tecnología de grupos.



tes, los cuales podrían agruparse en más o menos 50 familias de partes. Algunas ventajas al usar el concepto de tecnología de grupos son

1. *Inventarios reducidos de trabajo en proceso*. Cada célula de manufactura opera como una unidad independiente, permitiendo controles mucho más rígidos en el flujo de producción. No se necesitan grandes inventarios de trabajo en proceso para conservar tiempos de ciclo cortos. Un beneficio colateral de esto sería la reducción de colas en partes y la confusión que de ello resulta.

2. *Tiempos reducidos de montaje*. Como las células de manufactura se organizan según los tipos de partes, los montajes de las máquinas no deberían variar significativamente cuando se cambia de una parte a otra. Esto permite que las células operen con mucha mayor eficiencia.
3. *Costos reducidos del manejo de material*. Para una compañía que produce 10 000 partes, la distribución del proceso requeriría una vertiginosa variedad de rutas de partes. Si los volúmenes son grandes, los centros de procesos necesariamente tendrían que estar separados por grandes distancias, provocando así costos sustanciales en el manejo de materiales. Una distribución basada en tecnología de grupos resolvería este problema.
4. *Mejor programación*. Al aislar los agrupamientos de las partes, es mucho más fácil rastrear el flujo de producción en cada célula. Reducir los tiempos de ciclo y las colas del trabajo en proceso conduce a programas de entrega más confiables.

Existen varias desventajas en el enfoque de tecnología de grupos. Una es que puede ser difícil determinar las familias de partes adecuadas. Las partes pueden agruparse por tamaño y forma o por requerimientos del proceso de manufactura. El primer enfoque resulta más fácil pero no es tan efectivo para desarrollar las distribuciones. La forma en que se agrupan las partes en gran medida es una función del sistema de codificación que se usa para clasificar esas partes. (Groover y Zimmers, 1984, estudian varios sistemas de codificación de partes y cómo se relacionan con el concepto de tecnología de grupo). El agrupamiento por familias de partes según el flujo de manufactura requiere analizar con cuidado el flujo de producción (Burbidge, 1975). No obstante, es probable que este método no juegue un papel importante en una compañía con un gran número de partes.

Las distribuciones por tecnología de grupos pueden requerir duplicar algunas máquinas. Para que una célula de manufactura sea autocontenida, la célula debe tener todas las máquinas necesarias para generar el producto. Duplicar las máquinas podría ser caro y quizás conduciría a mayores tiempos totales de inactividad.

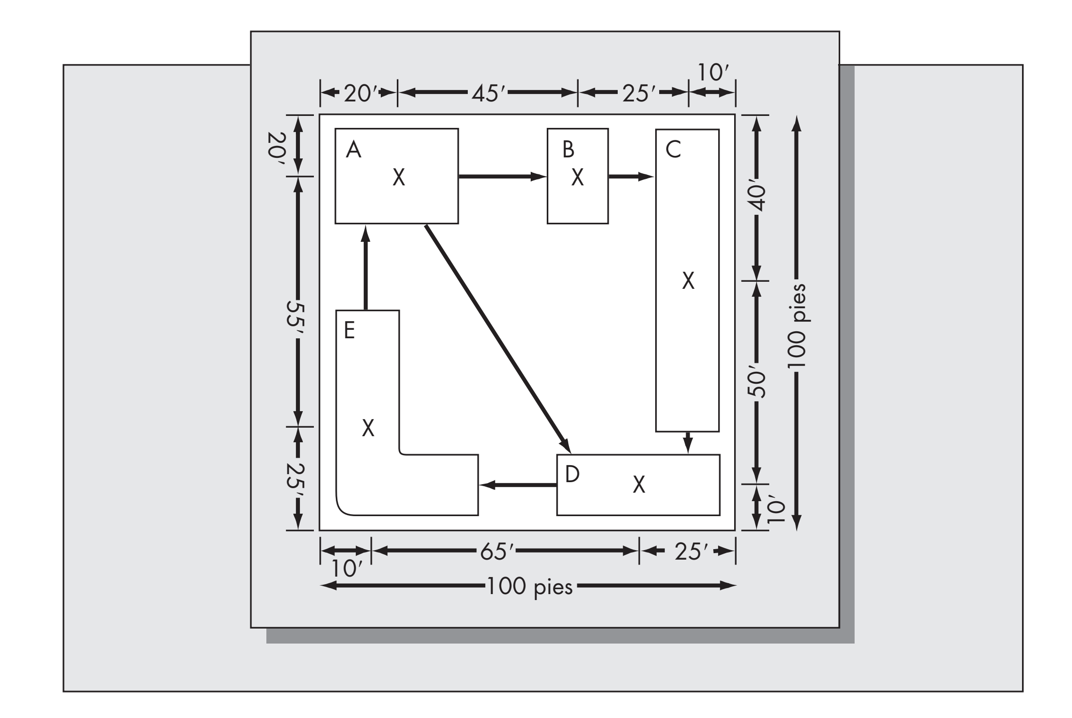
¿Bajo qué circunstancias sería preferible una distribución por tecnología de grupos a una distribución de proceso o producto puro? Un estudio de simulación de Sassani (1990) proporciona algunas respuestas. Sassani construyó una simulación de cinco células de manufactura. Inicialmente, si los productos de cada célula estaban bien definidos y las células se aislaban, el sistema funcionaba perfectamente. Sin embargo, cuando la mezcla del producto, su diseño respectivo y los patrones de demanda comenzaban a cambiar, la simulación mostraba que la eficiencia de la distribución disminuía.

La gran mayoría de las distribuciones existentes son de proceso o de producto. Las compañías que producen una amplia variedad de partes pueden escoger varias distribuciones para diferentes líneas de productos, o algún enfoque híbrido. La variedad del producto y el volumen anual son los principales factores que determinan la decisión apropiada. El enfoque de tecnología de grupos es relativamente nuevo. Está ganando aceptación lentamente porque las compañías combaten problemas tan difíciles como determinar apropiadamente los agrupamientos de las partes y cómo diseñar las células. Sin embargo, a medida que las fábricas automatizadas se tornan familiares, el concepto de tecnología de grupos podría jugar un papel más importante en la distribución de las plantas y en el diseño de proceso.

Problemas para las secciones 10.1 a 10.3

1. Para cada uno de los ocho objetivos de un estudio de distribuciones enlistado en la sección 10.1, proporcione casos donde dicho objetivo sea importante y ejemplos en los cuales no sea relevante.
2. Una instalación de manufactura consta de cinco departamentos: A, B, C, D y E. Para este momento, la instalación está dispuesta como en la figura 10-10. Desarrolle un diagrama dehacia sobre esta distribución. Estime la distancia que separa a los departamentos con base en el patrón de flujo del sistema de manejo de material de dicha figura. Suponga que todos los departamentos se ubican en sus centros como lo marca la figura 10-10.

FIGURA 10-10
Distribución de la instalación de manufactura (para el problema 2).



3. Cuatro productos se producen en la instalación descrita en el problema 2. Las rutas para estos productos y sus tasas de producción pronosticadas son

Producto	Rutas	Producción pronosticada (unidades/semana)
1	A–B–C–D–E	200
2	A–D–E	900
3	A–B–C–E	400
4	A–C–D–E	650

Suponga que los cuatro productos se producen en lotes de tamaño 50.

a) Convierta esta información a un diagrama de-hacia similar a la figura 10-4 pero presentando el número semanal de viajes de manejo de material entre departamentos.

b) Suponga que se estima que el costo de transporte de bienes por pie es de $2. Usando el diagrama de-hacia que determinó en el inciso *a*, obtenga un diagrama de-hacia que proporcione el costo del manejo de material por semana entre dos departamentos.

c) Desarrolle un diagrama de relación de actividades para esta instalación, con base en el diagrama de-hacia del costo del manejo de material que construyó en el inciso *b*. Use solamente las claves A, E, I, O, U, con A que se asigna al costo más elevado y U al más bajo.

d) Con base en los resultados de los incisos *b* y *c*, ¿recomendaría distribuir esta instalación de otro modo?

4. Considere el ejemplo del taller de máquinas con los diagramas de-hacia dados en las figuras 10-3 a 10-5. Con base en los resultados de la figura 10-5, ¿de cuántas maneras podría mejorarse la presente distribución?

5. Sugiera una distribución para el restaurante de comida rápida Meat Me, según el diagrama de relación mostrado en la figura 10-2. Suponga que la instalación tiene 50 pies cuadrados y la mitad del restaurante se destina a asientos para los clientes.
6. Describa las diferencias entre las distribuciones de producto, de proceso, y de tecnología de grupos. Describa las circunstancias en las cuales resultaría más apropiado cada tipo de distribución.

10.4 UN PROBLEMA PROTOTIPO DE DISTRIBUCIÓN Y EL MODELO DE ASIGNACIÓN

Se han desarrollado varias técnicas analíticas para ayudar en los problemas de distribución. Sin embargo, la mayoría de los problemas reales son demasiado complejos para resolverse con estos métodos sin una computadora. (Posteriormente estudiaremos el papel de la computadora en la distribución de las instalaciones en la sección 10.6.) Sin embargo, en algunos casos, los problemas de distribución pueden formularse como problemas de *asignación*, según se describe en este apartado. En especial, el modelo de asignación funciona cuando sólo consideramos un conjunto discreto de ubicaciones alternativas y los departamentos no interactúan. Aunque los problemas de asignación de tamaño moderado pueden resolverse a mano, los problemas *muy* grandes sólo pueden resolverse con ayuda de las computadoras.

Ejemplo *10.3*

Tras un incremento en el volumen de ventas, Sam Melder (el dueño de Melder, Inc., una pequeña compañía manufacturera ubicada en Muncie, Indiana) ha decidido expandir la capacidad de producción. Añadió una nueva ala a la planta, que va a alojar a cuatro máquinas: (1) una troqueladora, (2) una esmeriladora, (3) un torno, y (4) una máquina de soldadura. Existen solamente cuatro ubicaciones posibles para estas máquinas, digamos A, B, C, y D. Sin embargo, la máquina de soldadura, que es la máquina más grande, no va a caber en la ubicación B.

El superintendente de la planta ha estimado el impacto en términos de los costos de manejo de material de la ubicación de cada una de las máquinas en cada una de las ubicaciones posibles. Estos costos, expresados en términos de dólares por hora, se representan en la siguiente tabla.

		Ubicación			
		A	**B**	**C**	**D**
Máquinas	**1**	94	13	62	71
	2	62	19	84	96
	3	75	88	18	80
	4	11	*M*	81	21

La entrada *M* representa un costo muy grande. Se usa para indicar que la máquina 4 no se permite en la ubicación B. Como el objetivo consiste en distribuir las cuatro máquinas según las ubicaciones para reducir el costo total de manejo de material, una solución óptima nunca asignará la máquina 4 a la ubicación B.

Las asignaciones de costos mínimos rara vez pueden encontrarse por inspección. Por ejemplo, un enfoque razonable podría ser referir secuencialmente las máquinas a las ubicaciones de costo más bajo. El costo más bajo en la matriz es 11, de modo que la máquina 4 se ubicaría en la posición A. Si se eliminan el último renglón y la primera columna, vemos que el siguiente costo más bajo restante es 13, de modo que la máquina 1 se asignaría a la posición B. Si ahora eliminamos también el primer renglón y la segunda columna, el costo más pequeño restante es 18, de modo que la máquina 3 se asignaría a la posición C. Finalmente, la máquina 2 debe asignarse a la posición D. El costo total de esta solución es 11 + 13 + 18 + 96 = $138. Según veremos, esta solución es subóptima. (El costo óptimo es $114. ¿Puede usted encontrar la solución óptima?)

Un enfoque tan simple como éste rara vez conducirá a una solución óptima. Se presentará un algoritmo de solución a partir del cual puede obtenerse manualmente una solución óptima al problema de asignaciones para problemas de tamaño moderado.

El algoritmo de asignación

El algoritmo de solución que ahora presentamos se basa en la siguiente observación:

> Resultado: si se suma o se resta una constante de todas las entradas en un renglón o en una columna de una matriz de asignación de costos, la asignación óptima permanece sin cambio.

Al aplicar este resultado, el objetivo es continuar la resta de constantes de renglones y de columnas en la matriz de costos hasta que pueda asignarse costo cero. La asignación de costo cero para la matriz modificada es óptima para el problema original. Esto conduce al siguiente algoritmo para resolver problemas de asignación:

Procedimiento de solución para los problemas de asignación

1. Localice el número más pequeño en el renglón 1 y réstelo de todas las entradas del renglón 1. Repita para todos los renglones de la matriz de costos.
2. Localice el número más pequeño de la columna 1 y réstelo de todas las entradas de la columna 1. Repita para todas las columnas de la matriz de costos.
3. En este punto cada renglón y cada columna tendrá cuando menos un cero. Si es posible hacer una asignación de costo cero, entonces hágala. Esa será la solución óptima. Si no, vaya al paso 4.
4. Determine el número máximo de asignaciones al costo cero. Esto equivaldrá al número más pequeño de renglones necesarios para cubrir todos los ceros. Los renglones se encuentran por inspección y no son necesariamente únicos. Lo que importa es que el número de renglones extraídos no supere al número máximo de asignaciones de costo cero.
5. Escoja el número no cubierto más pequeño y haga lo siguiente:

   a) Réstelo de todos los otros números no cubiertos.

   b) Súmelo a los números donde se crucen los renglones.

   c) Regrese al paso 3.

El proceso continúa hasta que pueda hacerse una asignación de costo cero. Observe que el paso 5 es sólo una aplicación del resultado de la manera siguiente: si el número no cubierto más pequeño se resta de cada uno de los elementos de la matriz y luego se suma a cada elemento cubierto, se restará una vez y se sumará dos veces donde se crucen los renglones.

Ejemplo *10.3* *(continuación)*

Regresemos al ejemplo 10.3. La matriz original de costos es

		Ubicación			
		A	**B**	**C**	**D**
Máquinas	**1**	94	13	62	71
	2	62	19	84	96
	3	75	88	18	80
	4	11	*M*	81	21

Paso 1. Al restar el número más pequeño de cada renglón se obtiene

81	0	49	58
43	0	65	77
57	70	0	62
0	*M*	70	10

Como *M* es muy grande en relación con los otros costos, al restar 11 de *M* todavía se tiene un número muy grande, el cual nuevamente denotamos como *M* por comodidad. En este punto aparece cuando menos un cero en cada renglón, y en cada columna con excepción de la última.

Paso 2. Al restar 10 de cada número en la columna final se obtiene

81	0	49	48
43	0	65	67
57	70	0	52
0	*M*	70	0

En este punto tenemos cuando menos un cero en cada renglón y en cada columna (paso 3). Sin embargo, eso no significa necesariamente que sea posible asignar un costo cero. De hecho, es posible hacer cuando mucho tres asignaciones de costo cero en esta etapa, como se muestra en el paso 4.

Paso 4.

81	[0]	49	48
43	0	65	67
57	70	[0]	52
[0]	*M*	70	0

Las tres asignaciones mostradas en el paso 4 son 1-B, 3-C, y 4-A. También hay otras maneras de asignar tres ubicaciones a costo cero. No importa lo que escojamos en esta etapa, sólo sabemos que tres son posibles. El siguiente paso sería encontrar tres renglones que cubran todos los ceros. Esto se muestra a continuación.

81	0	49	48
43	0	65	67
~~57~~	~~70~~	~~0~~	~~52~~
~~0~~	~~*M*~~	~~70~~	~~0~~

Nuevamente, la selección de renglones no es única. Sin embargo, es importante no usar más de tres renglones. La selección de los tres renglones que cubran todos los ceros se hace por prueba y error.

Paso 5. El número no cubierto más pequeño, 43, se resta de todos los otros números no cubiertos y se suma a los números donde se cruzan los renglones. La matriz resultante es

		Ubicación			
		A	**B**	**C**	**D**
Máquinas	**1**	38	[0]	6	5
	2	[0]	0	22	24
	3	57	113	[0]	52
	4	0	*M*	70	[0]

Ahora es posible hacer una asignación de costo cero, como se muestra en la matriz. La asignación óptima es la máquina 1, la troqueladora, a la posición B; la máquina 2, la esmeriladora, a la posición A,

la máquina 3, el torno, a la posición C; la máquina 4, la soldadora, a la posición D. El costo total del manejo de material por hora, para la solución óptima, se obtiene recordando la matriz original de asignación de costos. Es 13 + 62 + 18 + 21 = $114 por hora.

El algoritmo de asignación también puede usarse cuando el número de sitios es mayor que el número de máquinas. Por ejemplo, suponga que haya seis sitios potenciales para ubicar las cuatro máquinas y que la matriz original de costos es

		Ubicación					
		A	**B**	**C**	**D**	**E**	**F**
Máquinas	**1**	94	13	62	71	82	25
	2	62	19	84	96	24	29
	3	75	88	18	80	16	78
	4	1 1	*M*	81	21	45	14

El procedimiento es agregar dos máquinas ficticias, 5 y 6, con costos cero. Entonces el problema se resuelve empleando el algoritmo de asignación como si hubiera seis máquinas y seis posiciones. Las posiciones de asignación para las máquinas ficticias son las que no se usan.

Problemas para la sección 10.4

7. Resuelva el siguiente problema de asignación:

	A	**B**	**C**	**D**
1	21	24	26	23
2	29	27	30	29
3	24	25	34	27
4	28	26	28	25

8. Se va a asignar cada una de cuatro máquinas a una de cinco posiciones posibles. El objetivo es asignar las máquinas a las posiciones que minimicen el costo del manejo de material. Los costos de máquina-posición se dan en la siguiente matriz. Encuentre la asignación óptima.

	A	**B**	**C**	**D**	**E**
1	26	20	22	21	25
2	35	31	33	40	26
3	15	18	23	16	25
4	31	34	33	30	*M*

9. La Universidad del Atlántico está mudando su escuela de negocios a un nuevo edificio, diseñado para alojar seis departamentos académicos. El tiempo promedio requerido por un estudiante para ir y venir de las clases en el edificio varía según la ubicación del departamento en el cual tome las clases. Con base en la distribución de las cargas de clase, el director ha estimado los siguientes tiempos medios de viaje por estudiante en minutos, dadas las ubicaciones departamentales.

		Ubicación					
		A	**B**	**C**	**D**	**E**	**F**
Departamento	**1**	13	18	12	20	13	13
	2	18	17	12	19	17	16
	3	16	14	12	17	15	19
	4	18	14	12	13	15	12
	5	19	20	16	19	20	19
	6	22	23	17	24	28	25

Encuentre la asignación óptima de los departamentos a las ubicaciones para minimizar el tiempo medio de viaje por cada estudiante dentro y fuera del edificio.

*10.5 FORMULACIONES MATEMÁTICAS MÁS AVANZADAS DE PROGRAMACIÓN

El modelo de asignación descrito en la sección 10.4 puede ser útil para determinar distribuciones óptimas en un número limitado de problemas reales. La limitación primaria del modelo de asignación simple es que, en la mayoría de los casos, el número de viajes de manejo de material y el costo asociado con el mismo se suponen independientes de la ubicación de las demás instalaciones. En el ejemplo 10.3, el costo de asignación de la troqueladora a la posición A se supuso de $94 por hora. Sin embargo, en la mayoría de los casos este costo también dependería de la posición de las otras máquinas.

Una formulación del problema que toma en cuenta esta característica resulta considerablemente más compleja. Para evitar la confusión, supondremos que el problema es asignar las máquinas a las posiciones. Por supuesto que el problema podría ser la asignación de otros tipos de subinstalaciones a localidades (específicas), pero retendremos esta terminología por comodidad. Defina las siguientes cantidades:

n = Número de máquinas;

c_{ij} = Costo por periodo de tiempo de asignación de la máquina i a la ubicación j. Este costo podría ser un costo de reubicación por una sola vez que se convierte a un equivalente anual;

d_{jr} = Costo de realizar un solo viaje de manejo de material de la ubicación j a la ubicación r;

f_{ik} = Número promedio de viajes por lapso de tiempo de la máquina i a la máquina k;

S_i = El conjunto de ubicaciones factibles de asignación para la máquina i;

$$a_{ijkr} = \begin{cases} f_{ik}d_{jr} & \text{si } i \neq k \text{ o } j \neq r, \\ c_{ij} & \text{si } i = k \text{ y } j = r; \end{cases}$$

$$x_{ij} = \begin{cases} 1 & \text{si la máquina } i \text{ se asigna a la ubicación } j, \\ 0 & \text{en otro caso.} \end{cases}$$

Interprete a_{ijkr} como el costo de manejo de material por unidad de tiempo si la máquina i se asigna a la ubicación j y la máquina k se asigna a la ubicación r. Se incurre en este costo únicamente si tanto x_{ij} como x_{kr} son iguales a uno. Así, se concluye que el costo total de la asignación de las máquinas a las ubicaciones está dado por

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{r=1}^{n} a_{ijkr}x_{ij}x_{kr}. \qquad (1)$$

Como se suman todos los índices de 1 hasta n, cada asignación se contará dos veces; de ahí la necesidad de multiplicar por $\frac{1}{2}$. Se incluyen restricciones para asegurar que cada máquina se asigne exactamente a una ubicación y cada ubicación se asigne exactamente a una máquina. Éstas son

$$\sum_{i=1}^{n} x_{ij} = 1, \qquad j = 1, \ldots, n; \tag{2}$$

$$\sum_{j=1}^{n} x_{ij} = 1, \qquad i = 1, \ldots, n; \tag{3}$$

$$x_{ij} = 0 \text{ o } 1, \qquad i = 1, \ldots, n \text{ y } j = 1, \ldots, n; \tag{4}$$

$$x_{ij} = 0, \qquad i = 1, \ldots, n \text{ y } j \notin S(i). \tag{5}$$

La formulación matemática de programación es minimizar el costo total de manejo del material de la asignación, (1), sujeto a las restricciones (2), (3), (4), y (5). Esta formulación se conoce como el problema de asignación cuadrática. En general, estos problemas son muy difíciles de resolver. Debe considerarse usar un método de este tipo solamente para problemas de tamaño moderado.

Problema para la sección 10.5

10. Considere el siguiente problema con dos ubicaciones y tres máquinas. Suponga que los costos para transportar una carga unitaria de la ubicación j a la ubicación r se dan en la siguiente tabla:

		A la ubicación	
		A	**B**
De la ubicación	**A**		6
	B	9	

El número de viajes promedio que se requieren de la máquina i a la máquina k por hora son

		A la máquina		
		1	**2**	**3**
	1	0	3	1
De la máquina	**2**	0	0	3
	3	0	4	0

No se toman en cuenta los costos de reubicación. Escriba la formulación completa de asignación cuadrática para este problema de ubicación.

10.6 TÉCNICAS COMPUTARIZADAS DE DISTRIBUCIÓN

La formulación de asignación cuadrática que se explicó en la sección 10.5 tiene varias insuficiencias. Primero, deben especificarse los costos de todos los viajes de manejo de material y el número esperado de viajes desde cada uno de los departamentos hacia los otros. Si intervienen muchos departamentos, podría ser difícil obtener esta información. Además, todavía no se ha descubierto una técnica eficiente para resolver los problemas grandes de asignación cuadrática.

Por estas razones ha habido un considerable interés en los métodos asistidos por computadora. Estos métodos son heurísticos; no garantizan una solución óptima pero generalmente dan soluciones eficientes. Pueden usarse para solucionar problemas demasiado grandes como para resolverse analíticamente. Los métodos que estudiamos en esta sección son CRAFT, COFAD, ALDEP, CORELAP, y PLANET.

En cada uno de estos métodos el objetivo es minimizar el costo de manejo de material. Sin embargo, las soluciones generadas por estos programas de computadora deben considerarse en el contexto del problema. No se consideran aspectos como la seguridad de la planta, el ruido y la estética. La distribución que se obtiene con un programa de computadora deberá modificarse tomando en cuenta estos factores.

Los programas de computadora para determinar distribuciones generalmente se dividen en dos clases: 1) las rutinas de mejora y 2) las rutinas de construcción. Una rutina de mejora toma una distribución existente y considera el efecto de intercambiar la ubicación de dos o más instalaciones. Una rutina de construcción construye la distribución desde cero proveniente de los datos de flujo y de la información del diagrama de relación de actividades. CRAFT y COFAD son ambas rutinas de mejora, y PLANET, CORELAP, Y ALDEP son rutinas de construcción. Las rutinas de mejora tienen la desventaja de requerir que se especifique una distribución inicial. Sin embargo, las rutinas de mejora generalmente conducen a distribuciones más prácticas. Las rutinas de construcción frecuentemente proporcionan distribuciones que contemplan departamentos con formas caprichosas.

CRAFT

CRAFT (técnica computarizada de asignación relativa de instalaciones, en inglés *computerized relative allocation of facilities technique*) fue una de las primeras rutinas de distribuciones asistidas por computadora que se han desarrollado. Como se observó anteriormente, CRAFT es una rutina de mejora y requiere que el usuario especifique una distribución inicial. El objetivo de CRAFT es minimizar el costo total de transporte de una distribución, donde el costo de transporte se define como el producto del costo de movimiento de una carga unitaria del departamento i al departamento j por la distancia entre los departamentos i y j. Para ser más específico, defina

n = Número de departamentos,

v_{ij} = Número de cargas que se mueven del departamento i al departamento j para un lapso de tiempo dado,

u_{ij} = Costo de mover una carga unitaria una distancia unitaria del departamento i al departamento j,

d_{ij} = Distancia que separa a los departamentos i y j.

Se concluye que $y_{ij} = u_{ij}v_{ij}$ es el costo de movimiento del flujo total del producto durante el intervalo de tiempo especificado para una distancia unitaria de i a j, y que $y_{ij}d_{ij}$ es el costo del flujo del producto de i a j. El costo total del flujo de producto entre todos los pares de departamentos es

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_{ij}d_{ij}.$$

Observe que la información de entrada v_{ij} y u_{ij} puede representarse como un diagrama dehacia, como se muestra en las figuras 10-3 y 10-4.

Dos supuestos implícitos que se hacen en CRAFT son

1. Los costos de movimiento son independientes de la utilización del equipo.
2. Los costos de movimiento son una función lineal de la longitud del movimiento.

Cuando estos supuestos no pueden justificarse, es posible usar CRAFT para minimizar el producto de los flujos por las distancias solamente mediante la asignación de un valor de 1 al costo unitario u_{ij}. Las entradas d_{ij} se calculan con el programa a partir de la especificación de una distribución inicial. Con base en una distribución inicial, CRAFT considera el intercambio de la posición de los departamentos adyacentes y calcula el costo de manejo de material en el

intercambio resultante. El programa escoge el intercambio por pares que conduce a la mayor reducción de costos.

Se supone que los departamentos tienen forma rectangular o están compuestos por piezas rectangulares. Además, se supone que los departamentos están ubicados en sus *centroides*. El centroide es un sinónimo de las coordenadas del centro de gravedad o del punto de equilibrio. En el apéndice 10-A se presenta un comentario sobre los centroides y cómo se calculan para objetos en el plano. La exactitud de la hipótesis de que un departamento se ubique en su centroide depende de la forma del departamento. La hipótesis es muy exacta si la forma del departamento es cuadrada o rectangular, pero es menos exacta para departamentos con formas complejas.

Para entender mejor cómo CRAFT determina las distribuciones, considere el siguiente ejemplo.

Ejemplo *10.4*

Hace poco, una compañía manufacturera local terminó la construcción de una nueva ala de un edificio existente para alojar cuatro departamentos: A, B, C y D. El ala mide 100 pies por 50 pies. El gerente de la planta ha escogido una distribución inicial para los cuatro departamentos. Esta distribución aparece en la figura 10-11. Hemos marcado la posición de los centroides de los departamentos con un punto. En la figura vemos que el departamento A requiere 1 800 pies cuadrados, B requiere 1 200 pies cuadrados, C requiere 800 pies cuadrados, y D requiere 1 200 pies cuadrados.

Una de las entradas que requiere CRAFT son los datos de flujo, es decir, el número de viajes de manejo de material por unidad de tiempo desde cada uno de los departamentos hacia cada uno de los demás departamentos. Los datos se dan en el diagrama de-hacia que se muestra en la figura 10-12a. Se supone que la distancia entre departamentos es la distancia rectilínea entre las posiciones de los centroides. En la figura 10-11, vemos que las posiciones de los centroides de la distribución inicial son

$$(x_A, y_A) = (30, 35), \qquad (x_C, y_C) = (20, 10),$$
$$(x_B, y_B) = (80, 35), \qquad (x_D, y_D) = (70, 10).$$

La distancia rectilínea entre A y B, por ejemplo, está dada por la fórmula

$$|x_A - x_B| + |y_A - y_B| = |30 - 80| + |35 - 35| = 50.$$

Se calculan las distancias entre otros pares de departamentos de una manera similar. Los cálculos se resumen en el diagrama de-hacia de la figura 10-12b. Observe que hemos supuesto que la distancia desde el departamento *i* hasta el departamento *j* es la misma que la distancia del departamento *j* al departamento *i*. Como dijimos antes, esto no necesariamente es cierto si el material se transporta solamente en una dirección.

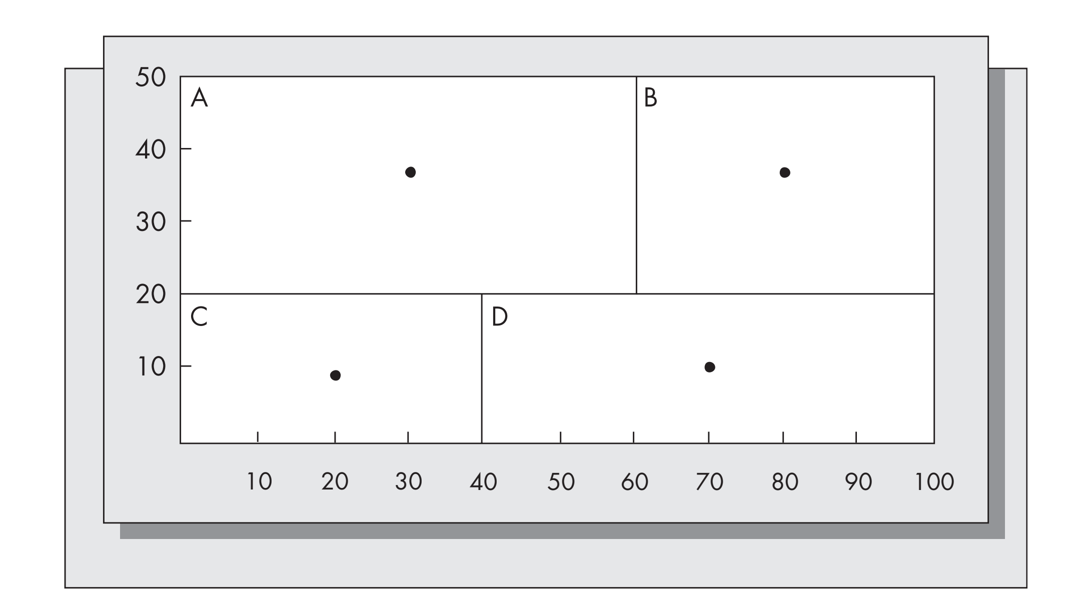
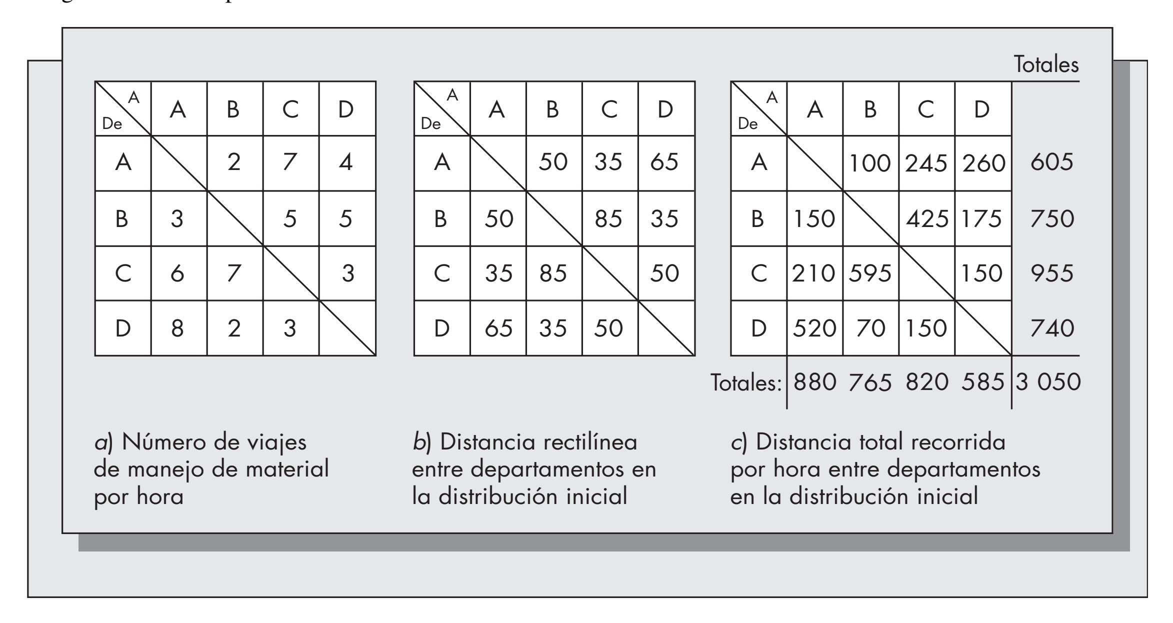


FIGURA 10-11 Distribución inicial para el ejemplo CRAFT (Ejemplo 10.4).

FIGURA 10-12
Diagramas de-hacia para la distribución inicial.



De \ A	A	B	C	D
A		2	7	4
B	3		5	5
C	6	7		3
D	8	2	3	

a) Número de viajes de manejo de material por hora

De \ A	A	B	C	D
A		50	35	65
B	50		85	35
C	35	85		50
D	65	35	50	

b) Distancia rectilínea entre departamentos en la distribución inicial

De \ A	A	B	C	D	Totales
A		100	245	260	605
B	150		425	175	750
C	210	595		150	955
D	520	70	150		740
Totales:	880	765	820	585	3 050

c) Distancia total recorrida por hora entre departamentos en la distribución inicial

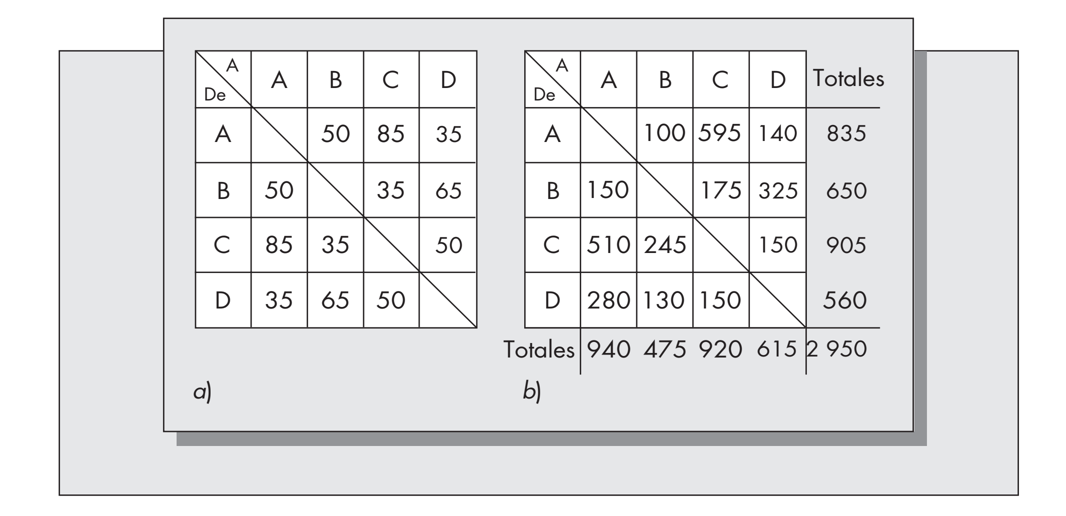
Normalmente, se requeriría un tercer diagrama de-hacia. Esto daría el costo de transporte de una carga unitaria para una distancia unitaria entre el departamento *i* y el departamento *j*. Como se observa, puede usarse CRAFT para minimizar el producto de los flujos por las distancias asignándoles valor de 1 a estos costos de transporte. Supondremos que éste es el caso en nuestro ejemplo. Así, se obtiene el costo horario de transporte de materiales hacia y desde cada uno de los departamentos simplemente multiplicando las entradas en el diagrama de-hacia de la figura 10-12*a* por las entradas en el diagrama de-hacia de la figura 10-12*b*. Estos cálculos se resumen en la figura 10-12*c*. La distancia total recorrida por hora para la distribución inicial es 3 050 pies.

Ahora CRAFT considera el efecto de los intercambios por pares de dos departamentos que tengan límites adyacentes o que tengan la misma área. (CRAFT también puede considerar el efecto de intercambio de las ubicaciones de tres departamentos. No consideremos esta opción en nuestro ejemplo.) El resultado del intercambio se determina intercambiando la ubicación de los centroides. Sin embargo ésta es solamente una aproximación, puesto que al intercambiar las ubicaciones de dos departamentos no necesariamente se intercambiará la ubicación de sus centroides.

Si tuviéramos que intercambiar las ubicaciones de A y de B, por ejemplo, supondríamos que $(x_A, y_A) = (80, 35)$ y $(x_B, y_B) = (30, 35)$, que conduciría al nuevo diagrama de distancias de-hacia que aparece en la figura 10-13*a*. Al multiplicar los datos originales de flujo de la figura 10-12*a* por este diagrama de distancias, se obtiene un nuevo diagrama de costos, que se muestra en la figura 10-13*b*. El intercambio de los centroides de A y B conduce a la reducción pronosticada de costos de 3 050 a 2 950, o sea aproximadamente 3 por ciento. (El costo real de intercambio de las ubicaciones de A y B sería un poco distinto porque los centroides no se intercambian exactamente.)

CRAFT considera todos los intercambios por pares de los departamentos adyacentes o de los departamentos con áreas idénticas y selecciona el que conduce al mayor decremento en el costo pronosticado. No presentaremos los detalles de los cálculos sino simplemente resumiremos los resultados. El intercambio de los centroides de A y C conduce a un costo total pronosticado de 2 715, y al intercambiar los centroides de A y D se obtiene un costo total pronosticado de 3 185. También deben considerarse otros dos intercambios: B y D, y C y D. Observe que no se considera el intercambio de B y C porque no tienen áreas iguales y no comparten un límite común. El intercambio de los centroides de B y D conduce a un costo total pronosticado de 2 735, y al intercambiar los centroides de C y D se obtiene un costo total pronosticado de 2 830.

FIGURA 10-13
Nuevos diagramas de-hacia para distancia y costo después de intercambiar los centroides para A y B.



De \ A	A	B	C	D
A		50	85	35
B	50		35	65
C	85	35		50
D	35	65	50	

a)

De \ A	A	B	C	D	Totales
A		100	595	140	835
B	150		175	325	650
C	510	245		150	905
D	280	130	150		560
Totales	940	475	920	615	2 950

b)

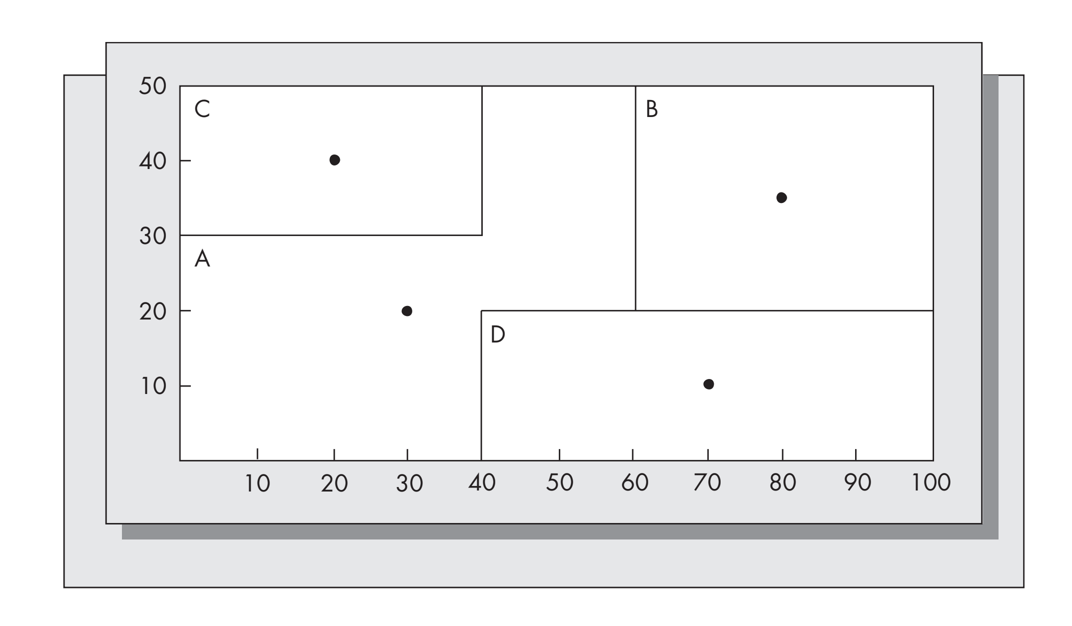
La reducción máxima pronosticada en el costo se alcanza intercambiando A y C. La nueva distribución con el intercambio de A y C se muestra en la figura 10-14. Como C tiene el área más pequeña, se coloca en la esquina superior izquierda del espacio anteriormente ocupado por A, de modo que el espacio restante permita que A sea contiguo. Observe que A ya no es rectangular. El costo real de la nueva distribución no es necesariamente igual al valor pronosticado de 2 715. El centroide de A se calcula con el uso del método esbozado en el apéndice 10-A. Se determina encontrando primero los momentos M_x y M_y dados por

$$M_x = (40^2 - 0)(30 - 0)/2 + (60^2 - 40^2)(50 - 20)/2 = 54000,$$
$$M_y = (30^2 - 0)(40 - 0)/2 + (50^2 - 20^2)(60 - 40)/2 = 39000,$$

y dividiendo entre el área de A para obtener

$$x_A = 54000/1800 = 30,$$
$$y_A = 39000/1800 = 21.66667.$$

FIGURA 10-14
Nueva distribución con A y C intercambiados.



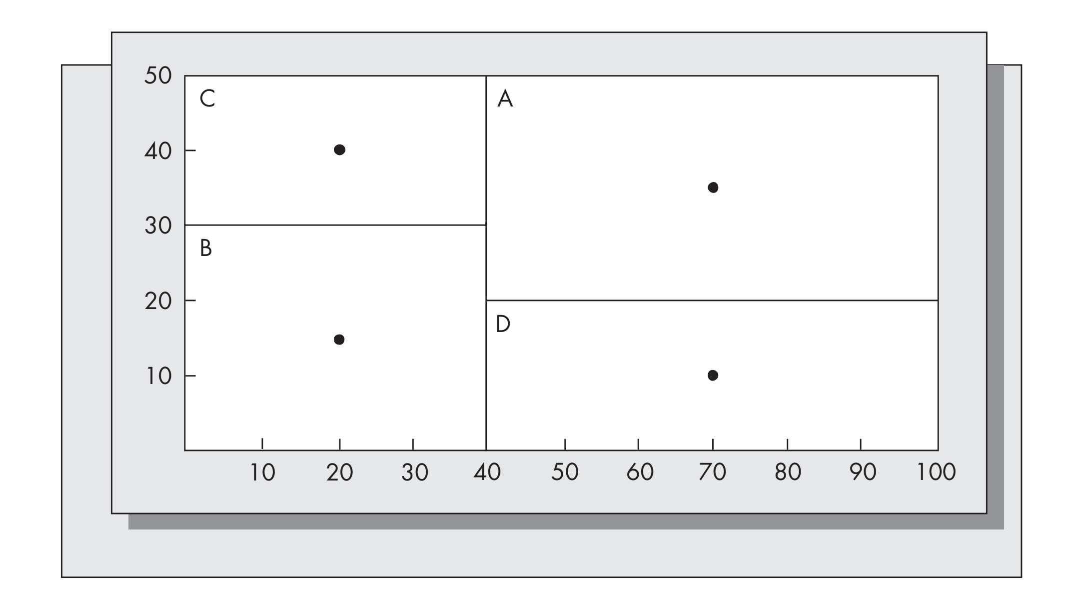


FIGURA 10-15
Segunda iteración, obtenida al intercambiar las ubicaciones de A y B.

El centroide de C se ubica en el centro de simetría, que es

$$(x_C, y_C) = (20, 40).$$

Los centroides de B y D permanecen sin cambio. En realidad el costo de la nueva distribución es 2 810, que es un poco mayor al valor de 2 715 pronosticado en el último paso, pero que es todavía menos que el costo original. El proceso continúa hasta que ya no sean posibles las reducciones de costo. En este punto solamente es necesario considerar tres intercambios: A y B, A y D, y B y D. Obviamente, no necesitamos considerar el intercambio de A y C, y C y D no pueden intercambiarse porque no comparten un borde común. El costo pronosticado que resulta del intercambio de los centroides de A y B es 2 763.33; de A y D, 3 641.33; y de B y D, 2 982. Es evidente que la mayor reducción pronosticada se alcanza ahora intercambiando las ubicaciones de A y B.

La nueva distribución se ilustra en la figura 10-15. Los centroides de los respectivos departamentos son ahora

$$(x_A, y_A) = (70, 35), \qquad (x_C, y_C) = (20, 40),$$
$$(x_B, y_B) = (20, 15), \qquad (x_D, y_D) = (70, 10).$$

El costo real de esta distribución es 2 530, el cual es considerablemente menor que el pronosticado en el paso anterior. El proceso continúa hasta que ya no puedan alcanzarse mayores reducciones en los costos pronosticados. Como A y B se intercambiaron en el paso anterior, no es necesario considerar ese intercambio nuevamente. El intercambio de A y C conduce a un costo pronosticado de 3 175; al intercambiar A y D, se obtiene un costo pronosticado de 2 753; y el intercambio de B y D, conduce a un costo pronosticado de 3 325. (No consideramos el intercambio de C y D en esta etapa porque no son contiguos y porque no tienen áreas iguales.) Como ninguno de los costos pronosticados es menor que el costo actual, terminamos los cálculos. La distribución recomendada por CRAFT, ilustrada en la figura 10-15, requirió dos iteraciones y condujo a una reducción de la distancia total recorrida desde 3 050 pies hasta 2 530 pies, o sea aproximadamente 17 por ciento.

COFAD

Como se observa, CRAFT es una rutina de mejora. Requiere que el usuario especifique una distribución inicial y que prosiga mejorando la distribución mediante intercambios por pares de departamentos contiguos. Una rutina de mejora similar a CRAFT es COFAD (diseño computarizado de instalaciones, en inglés *computerized facilities design*). En este capítulo no se presentarán los detalles de COFAD, sino que se revisan brevemente las mejoras que ofrece sobre CRAFT.

COFAD es una modificación de CRAFT que incorpora la selección de un sistema de manejo de material como parte del proceso de decisiones. Para una distribución dada, COFAD calcula el costo total de movimiento de la distribución para diferentes alternativas de manejo de material y selecciona la que tenga el menor costo. Se hacen mejoras considerando alternativas de equipo e intercambiando las asignaciones del equipo mal utilizado con equipo que tenga mejor utilización. Una vez que se determina la mejor selección de equipo para una distribución dada, el programa considera optimizar la distribución en una forma muy parecida a CRAFT. El programa termina cuando ya no son posibles mejoras adicionales. Una vez que el programa declara haber alcanzado un estado permanente, el problema puede resolverse nuevamente usando los diagramas de-hacia obtenidos al perturbar los datos originales desde 10 hasta 50 por ciento. El objetivo de volver a resolver el problema con los nuevos datos es asegurar que se ha alcanzado la solución óptima, y probar la sensibilidad de la solución a los datos de flujo.

ALDEP

Tanto CRAFT como COFAD son rutinas de mejora. Es decir, comienzan con una distribución inicial y consideran mejoras intercambiando las ubicaciones de los departamentos. La otra clase de técnicas computarizadas de distribución incluye programas que desarrollan las distribuciones esencialmente a partir de cero; no requieren de una distribución inicial. La experiencia ha demostrado que los programas de construcción frecuentemente tienden a dar como resultado distribuciones complejas, y por esa razón no han recibido tanta atención en la literatura como CRAFT.

ALDEP (programa automatizado para el diseño de distribuciones, en inglés *automated layout design program*) es una rutina de construcción más que una rutina de mejora. Esto implica que las distribuciones se determinan a partir de cero; el programa no requiere que el usuario especifique una distribución inicial. ALDEP emplea las clasificaciones de cercanía que se muestran en el diagrama de relación de actividades. La figura 10-2 es un ejemplo de un diagrama de relación. ALDEP selecciona primero un departamento en forma aleatoria y lo coloca en la esquina superior izquierda de la distribución. El siguiente paso consiste en explorar la gráfica de relación y colocar un departamento con una alta clasificación de cercanía (A o E) contiguo al primer departamento. Se colocan departamentos sucesivos en el área existente siguiendo un arreglo de arriba hacia abajo, de acuerdo con un patrón de “barrido”. El proceso continúa hasta que se han colocado todos los departamentos. En ese punto, se calcula una puntuación para la distribución. La puntuación se basa en una escala numérica asociada con las clasificaciones de cercanía. ALDEP usa los siguientes valores de clasificación de cercanía:

$$A = 4^3 = 64,$$
$$E = 4^2 = 16,$$
$$I = 4^1 = 4,$$
$$O = 4^0 = 1,$$
$$U = 0,$$
$$X = -4^5 = -1\,024.$$

El proceso completo se repite varias veces, y se escoge la distribución con la puntuación mayor. Ya que ALDEP procura alcanzar una alta puntuación en la clasificación de cercanía, frecuentemente recomienda distribuciones que tienen departamentos con formas muy fuera de lo común. El patrón de barrido ayuda a evitar este problema, pero implica que la distribución resultante siempre aparece como un conjunto de franjas contiguas. Podría ser que este tipo de distribución no sea apropiado ni conveniente para algunas aplicaciones. El usuario especifica el ancho de barrido que debe usarse; la selección del ancho de barrido puede tener un efecto significativo en la configuración de la distribución final. Por ejemplo, suponga que el departamento A tiene 14 cuadros, B tiene 8 cuadros, el ancho de barrido es de 2 cuadros, y el ancho de la

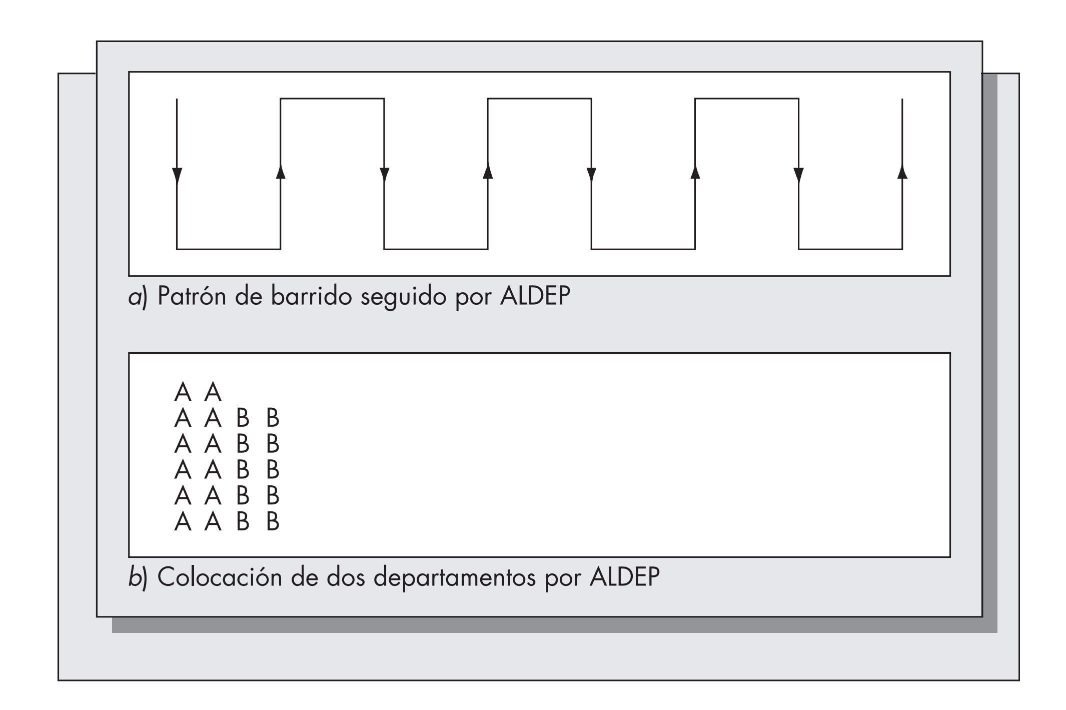


FIGURA 10-16
Método de colocación de departamentos usado por ALDEP.

instalación es de 6 cuadros. La distribución de estos dos departamentos dada por ALDEP se muestra en la figura 10-16*b*, suponiendo que los departamentos se colocan en el orden A, B.

CORELAP

Al igual que ALDEP, CORELAP (planeación computarizada de distribución de relaciones, en inglés *computerized relationship layout planning*) es una rutina de construcción que distribuye los departamentos usando claves de clasificación de cercanía. La principal diferencia entre CORELAP y ALDEP es que CORELAP no selecciona en forma aleatoria el primer departamento que se colocará. En vez de ello, se calcula una clasificación total de cercanía (TCR) para cada departamento. La TCR se basa en los valores numéricos $A = 6$, $E = 5$, $I = 4$, $O = 3$, $U = 2$, $X = 1$. Cada departamento se compara con los demás departamentos para obtener una TCR.

Por ejemplo, considere el restaurante de comida rápida Meat Me descrito en el ejemplo 10.1, con la gráfica de relación mostrada en la figura 10-2. Las claves de clasificación de cercanía para el departamento de cocinado de hamburguesas son *X, I, U, U,* y *U*, con lo que se obtiene una TCR para este departamento de 11. Los valores restantes de la TCR son

Cocinado de papas fritas:	*X, I, U, U, U*—11.
Empaque y almacenaje:	*I, I, O, E, E*—21,
Despachadores de refrescos:	*U, U, O, A, A*—19,
Servicio de mostrador:	*U, U, E, A, A*—21,
Servicio para automovilistas:	*O, A, E, U, U*—18.

CORELAP ahora selecciona al departamento con la mayor TCR y coloca a ese departamento en el centro de la instalación existente. Si hay un empate (como en este caso entre el departamento de empaque y almacenaje, y el departamento de servicio de mostrador), se coloca primero al departamento con la mayor área. Una vez seleccionado el departamento inicial, se explora el diagrama de relación para ver cuáles departamentos tienen la clasificación de cercanía mas alta en comparación con el departamento recién colocado. Suponga que el servicio de mostrador se coloca primero. Entonces, al explorar la figura 10-2, vemos que solamente el

departamento de despachadores de refrescos tiene un código de clasificación de cercanía de *A* en relación con el departamento de servicio de mostrador. Entonces, los despachadores de refrescos se colocarían a continuación en la distribución. Cada vez que un departamento sea elegido para colocarlo en la distribución, se consideran colocaciones alternativas y se calculan clasificaciones de ubicación basadas en valores numéricos especificados por el usuario para los códigos del diagrama de relación. Se coloca al nuevo departamento para maximizar la calificación de la colocación. (Pueden encontrarse ejemplos detallados de este procedimiento en Francis y White, 1974; y Tompkins y White, 1984).

PLANET

PLANET (técnica de análisis y evaluación de distribución de planta, en inglés *plant layout analysis and evaluation technique*), requiere esencialmente las mismas entradas que COFAD y, a diferencia de CORELAP y ALDEP, no usa información contenida en el diagrama de relación para generar una distribución. Sin embargo, a diferencia de CRAFT, PLANET no es una rutina de mejora. PLANET convierte los datos de entrada a un diagrama de flujo interno que proporciona el costo de mandar una unidad de flujo entre cada par de departamentos. Además de las entradas del diagrama de flujo interno, el programa también requiere que el usuario introduzca una clasificación de prioridades para cada departamento. La clasificación de prioridades es un número del 1 a 9, donde 1 representa la prioridad más alta. El programa selecciona departamentos para que entren en la distribución de manera secuencial, con base primero en la clasificación de prioridades y segundo en las entradas del diagrama de flujo interno. Como PLANET no restringe la distribución final para que se apegue la forma del edificio ni permite fijar la ubicación de ciertos departamentos, tiene los mismos problemas que otras rutinas de construcción. Las distribuciones obtenidas de PLANET, al igual que las de ALDEP y CORELAP, pueden conducir a departamentos con formas poco realistas. Por esa razón, estos métodos son mejores para darle al planeador algunas ideas alternas y distribuciones iniciales, más que para darle una solución final.

Métodos computarizados contra planificadores humanos

En la literatura sobre ciencias administrativas apareció un interesante debate relacionado con la efectividad de las técnicas computarizadas de distribución contra el juicio humano experto. El debate fue iniciado por un artículo de Scriabin y Vergin (1975). En su estudio, los autores compararon las distribuciones producidas por tres rutinas computarizadas de distribución, incluyendo CRAFT, y las distribuciones obtenidas sin el uso de estos programas específicos por 74 personas entrenadas en técnicas manuales de distribución. Los autores mostraron que la mejor solución obtenida por grupos de 20 de las 74 personas era mejor que la mejor solución obtenida por las tres rutinas de computadora. La diferencia más grande en el costo total ocurrió cuando el número de los departamentos era mayor (20). Las diferencias en porcentaje variaban de 0 a 6.7. Los autores concluyeron que por lo general los humanos superan a las rutinas de computación para las distribuciones porque "en problemas de mayor tamaño, la capacidad del hombre para reconocer y visualizar los patrones complejos le da una ventaja sobre los procedimientos esencialmente mecánicos que siguen los programas de computadora".

Las conclusiones de Scriabin y Vergin fueron atacadas por muchos investigadores. Buffa (1976) observó que el aspecto de la dominancia de flujo no se trató apropiadamente. La dominancia de flujo se refiere a la tendencia de los materiales a fluir en la misma secuencia en la distribución. La dominancia completa ocurre cuando una fábrica manufactura únicamente un solo producto de una manera idéntica cada vez. En un caso como éste, la distribución final puede obtenerse fácilmente mediante inspección visual. El otro extremo ocurre cuando los productos fluyen aleatoriamente a través de la fábrica. En ese caso todas las distribuciones son equivalentes. (Block, 1977, da una definición formal de dominancia de flujo.) Buffa afirmó que para los problemas en los cuales la dominancia de flujo fuera mayor que el 200 por ciento, podría esperarse que las personas obtuvieran fácilmente buenas soluciones. La dominancia de

flujo promedio en los problemas usados por Scriabin y Vergin fue evidentemente mayor que el 200 por ciento.

Otro problema con sus conclusiones, como lo señala Coleman (1977), fue que en un ambiente industrial típico no se tienen a 20 profesionistas resolviendo un problema de distribución. Al comparar las distribuciones producidas por los programas de computadora con aquellas producidas por las personas, Coleman mostró que las distribuciones generadas por computadora eran superiores. Finalmente, Block (1977) realizó un experimento aparte para probar la aseveración de Buffa de que cuando la dominancia de flujo estuviera por debajo de 200 por ciento, la computadora superaría al humano. Block comparó las distribuciones producidas por CRAFT con aquellas producidas por ocho humanos (cuatro ingenieros y cuatro técnicos en distribución) para un conjunto de problemas de distribución generados por Nugent, Vollmann y Ruml (1968), en los cuales la dominancia de flujo promedio era del 115 por ciento. Encontró que COFAD obtuvo mejores resultados en cada caso. Las mayores diferencias se observaron cuando el número de los departamentos era mayor que 10.

¿Qué podemos aprender de estos estudios? Parecería que las conclusiones de Scriabin y Vergin no se justifican. Resultados posteriores mostraron que los métodos asistidos por computadora son de hecho muy útiles y pueden conducir a distribuciones mucho mejores que las que pueden obtenerse por métodos simples visuales o gráficos cuando el número de los departamentos es grande y cuando los patrones de flujo material son muy variables.

Distribuciones dinámicas de planta

Hasta ahora, este capítulo ha considerado solamente problemas de distribución estática. Es decir, supusimos que los costos que aparecen en los diagramas de-hacia son fijos. Sin embargo, en algunas circunstancias, esta hipótesis podría no ser muy exacta. Nicol y Hollier (1983) observan que: “Frecuentemente ocurren cambios radicales en la distribución y por lo tanto la gerencia debería tomar esto en cuenta en su planeación a futuro”. Una distribución de planta dinámica se basa en reconocer que las demandas sobre el sistema, y por lo tanto los costos de cualquier configuración dada, pueden cambiar con el tiempo. Si se dispone de información sobre cómo cambiará el ambiente, la misma puede incorporarse a un modelo permitiendo que los costos en los diagramas de-hacia varíen para cada periodo de planeación. Este es precisamente el escenario considerado por Rosenblatt (1986), quien mostró cómo los diagramas de-hacia dinámicos formarían la base de un modelo de planeación de periodos múltiples y desarrolló un esquema de programación dinámica para resolver el sistema de ecuaciones resultante. Con un ejemplo, demuestra cómo las distribuciones dinámicas pueden conducir a ahorros en el costo para distribuciones fijas en un ambiente cambiante. Los modelos de este tipo cobrarán aún mayor importancia a medida que las compañías sigan moviéndose hacia estructuras basadas en una manufactura ágil y en mayor flexibilidad.

Otros métodos de computadora

Se han desarrollado otros métodos analíticos además de los ya estudiados. Un método llamado muestreo sesgado, sugerido por Nugent, Vollmann, y Ruml (1968), es una modificación directa de CRAFT que incluye un procedimiento de aleatorización para seleccionar los departamentos que van a intercambiarse. El método es mucho más lento computacionalmente, pero con frecuencia conduce a mejores distribuciones. Considerando los enormes progresos que se han logrado en la tecnología de computación, los aspectos computacionales son menos graves hoy de lo que eran en 1968. SPACECRAFT (véase Johnson, 1982) es una extensión de CRAFT diseñada con el fin de producir distribuciones para estructuras de niveles múltiples. Las dos modificaciones principales que se necesitaban eran que: 1) debían incorporarse restricciones en los procedimientos computacionales que permitieran que ciertos departamentos se ubicaran en pisos específicos, y 2) la no linealidad del tiempo requerida para transitar entre pisos precisaba cálculos de costo más complejos.

No todos los problemas de distribución son fácilmente accesibles a los métodos revisados aquí. Existen aplicaciones en las cuales no es posible suponer departamentos con forma

rectangular. Entre los ejemplos están las distribuciones de edificios de oficinas, el tablero de instrumentos de un avión, una ciudad o un vecindario, o un circuito integrado. Drezner (1980) introdujo un método llamado DISCON (dispersión y concentración, en inglés *dispersion and concentration*) que considera las instalaciones como discos. Primero los discos se dispersan matemáticamente en el plano, usando un sistema de ecuaciones diferenciales, las cuales simulan un sistema de resortes que conectan a los discos. La fase de dispersión proporciona una solución inicial, que posteriormente se mejora en la fase de concentración. El enfoque de Drezner trabaja mejor que CRAFT si los departamentos no son rectangulares y si el número de instalaciones es grande. También se han tomado en cuenta técnicas basadas en los métodos de teoría de gráficas (Foulds, 1983), y en el análisis estadístico de grupos (Scriabin y Vergin, 1985).

El software "clásico" revisado en esta sección ilustra los conceptos esenciales empleados en la mayoría de los productos de software para el diseño y la distribución de las instalaciones. Sin embargo, con la proliferación de la computadora personal en años recientes, en la actualidad se dispone de nuevas líneas de productos de software comercialmente disponibles para las computadoras basadas en Windows, Macintosh, y estaciones de trabajo UNIX. Los productos especiales han sido diseñados para segmentos del mercado bien definidos. El diseño de oficina, las líneas de proceso, y la planeación de instalaciones industriales son inherentemente problemas de distribución muy distintos, que requieren productos de software diseñados especialmente.

El problema del diseño de instalaciones de oficinas requiere formar bloques de espacio para actividades bien definidas. Estos problemas conciernen a los diseñadores de interiores y a los arquitectos. Las herramientas de dibujo con librerías extensas de símbolos son las más populares para estas aplicaciones. La más famosa es AutoCAD, que permite distribuciones en dos y tres dimensiones. Comúnmente las tuberías de procesos y el diseño de distribuciones se usan conjuntamente con las plantas químicas y casi siempre se despliegan como distribuciones tridimensionales. La distribución y el diseño de instalaciones industriales generalmente incluyen alguna variante de las herramientas de software estudiadas en este capítulo, o el uso de un paquete de simulación basado en gráficos especialmente diseñado, tal como ProModel. Sly (1995) lista aproximadamente 20 paquetes diferentes de software para la distribución y el diseño de instalaciones y describe muchas de sus características.

El diseño y la distribución inteligentes de las instalaciones industriales es un problema cuya importancia va en aumento. Por ejemplo, el costo de una nueva fábrica (instalación de fabricación) en la industria de los semiconductores comúnmente sobrepasa los mil millones de dólares. Se requieren inversiones similares para nuevas plantas de ensamblado automotriz. Las compañías no pueden darse el lujo de cometer errores al planear la organización de estas instalaciones, excesivamente caras.

Problemas para la sección 10.6

11. Describa brevemente cada una de las siguientes técnicas computarizadas de distribución. En cada caso, indique si el método es de construcción o de mejora.

    a) CRAFT
    b) COFAD
    c) ALDEP
    d) CORELAP
    e) PLANET

12. *a*) Discuta las ventajas y las desventajas al usar programas de computadora para desarrollar distribuciones.

    b) ¿Sugieren los resultados de los estudios analizados en la sección 10.6 que los planificadores humanos o los programas de computadora producen distribuciones superiores? ¿Por qué razones estos estudios no son concluyentes?

13. Para el ejemplo 10.4, donde se ilustró CRAFT, verifique los valores de los costos pronosticados (2 715, 3 185, 2 735 y 2 830) obtenidos en la primera iteración.

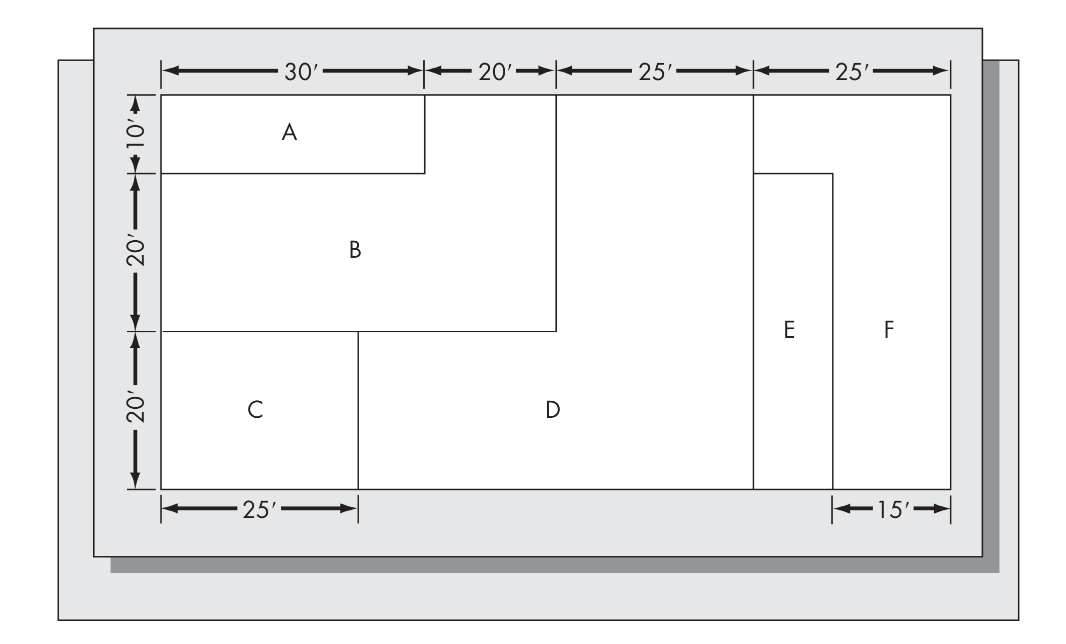
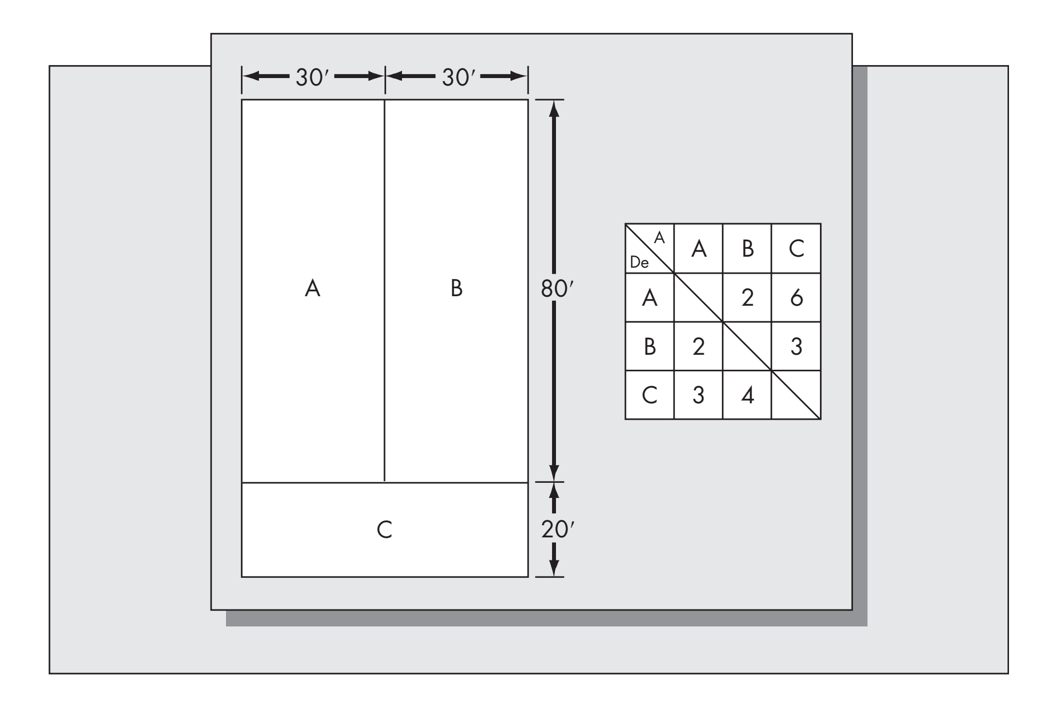


FIGURA 10-17
Distribución (para el problema 15).

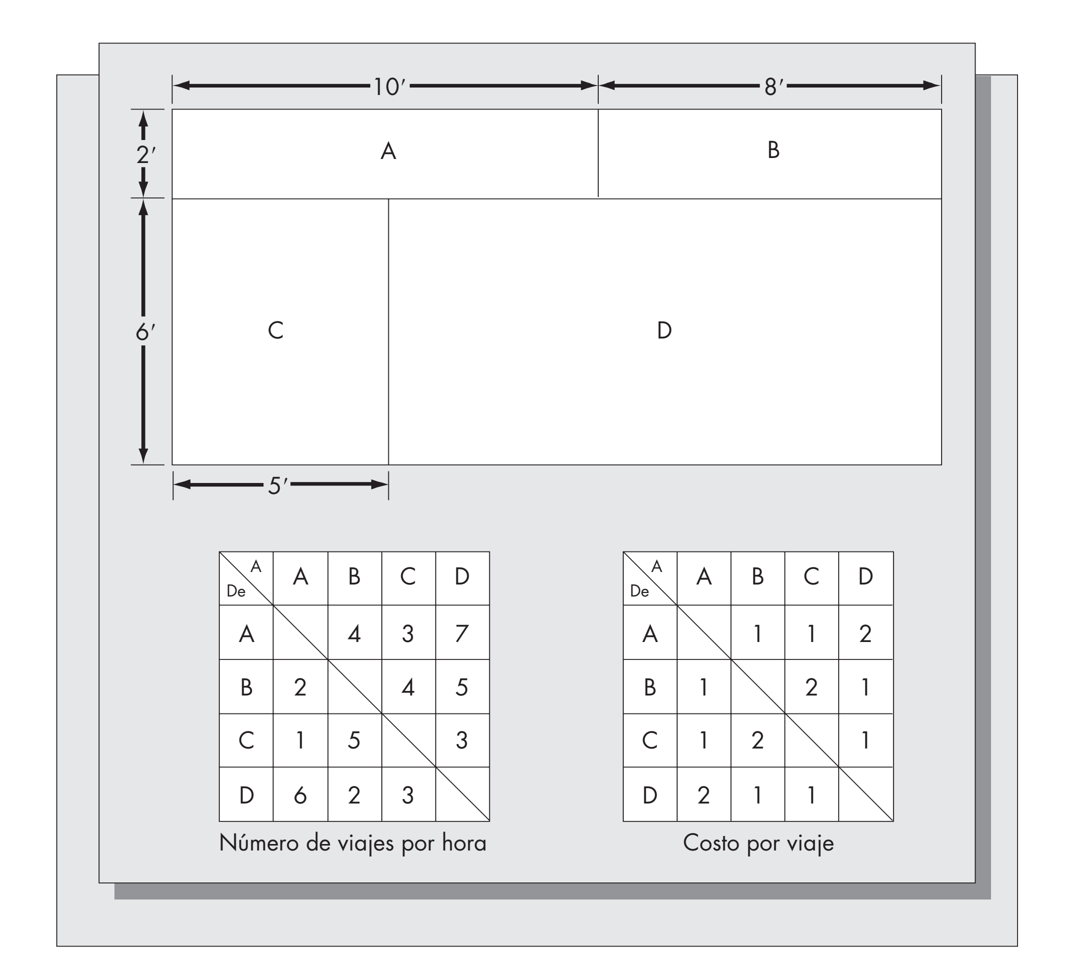


De \ A	A	B	C
A		2	6
B	2		3
C	3	4	

FIGURA 10-18
Distribución y diagrama de-hacia (para el problema 16).

14. Considere la distribución inicial para el ejemplo 10.4, que se muestra en la figura 10-11. Dibuje una figura que muestre la distribución obtenida después de intercambiar las ubicaciones de A y D, y encuentre los centroides de A y D en la nueva distribución.
15. Determine los centroides para los seis departamentos de la distribución ilustrada en la figura 10-17 usando los métodos esbozados en el apéndice 10-A.
16. Considere la distribución inicial y el diagrama de-hacia que se muestra en la figura 10-18. Suponiendo que la meta es minimizar la distancia total recorrida, determine la distribución final recomendada por CRAFT usando el método de intercambio por pares descrito en esta sección. Compare las cifras pronosticadas y las reales para la distancia total recorrida en la distribución final.

FIGURA 10-19
Distribución y diagramas de-hacia (para el problema 17).



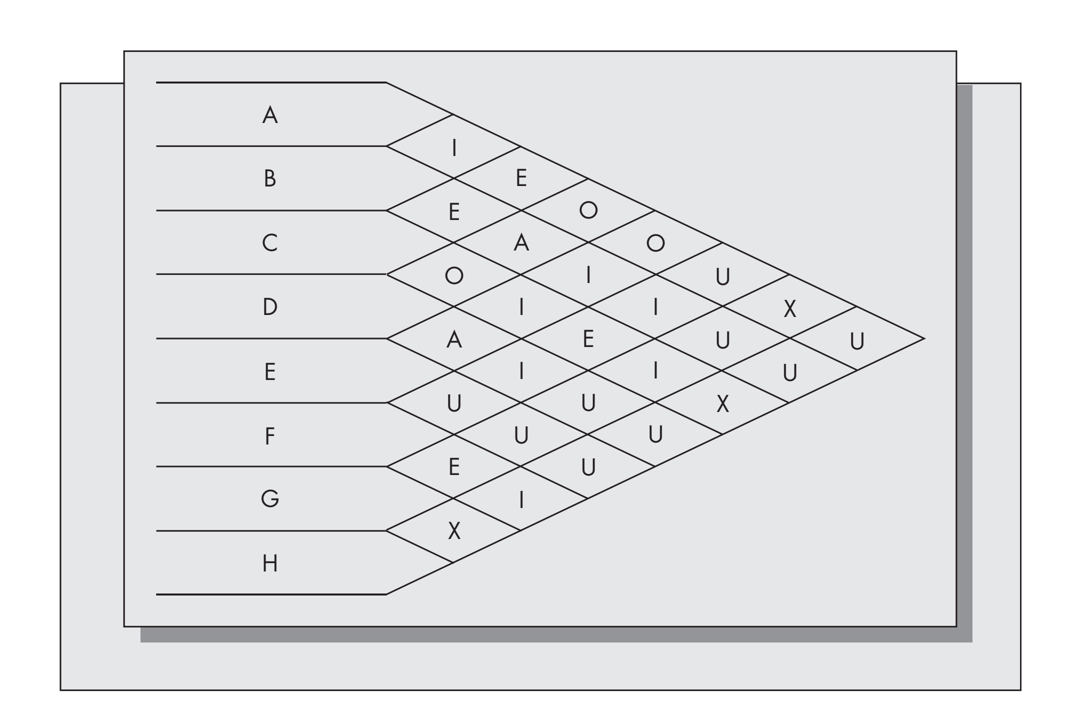
De \ A	A	B	C	D
A		4	3	7
B	2		4	5
C	1	5		3
D	6	2	3	

Número de viajes por hora

De \ A	A	B	C	D
A		1	1	2
B	1		2	1
C	1	2		1
D	2	1	1	

Costo por viaje

FIGURA 10-20
Gráfica de relación de actividades (para el problema 18).



17. Considere la distribución inicial ilustrada en la figura 10-19 y los dos diagramas de-hacia que muestran los datos de flujo y de costo. Use el enfoque CRAFT para obtener una distribución final. Compare el costo total de cada distribución pronosticada por CRAFT con el costo real de cada distribución.
18. Una instalación que consta de ocho departamentos tiene el diagrama de relación de actividades que se ilustra en la figura 10-20. Calcule la TCR para cada departamento que CORELAP usaría.
19. Seis departamentos, A, B, C, D, E, y F, consumen respectivamente 12, 6, 9, 18, 22, y 6 cuadros. Con base en un ancho de barrido de 4 y un tamaño de instalación de 5 por 16, use las técnicas empleadas por ALDEP para encontrar una distribución. Suponga que los departamentos están ordenados alfabéticamente.

10.7 SISTEMAS FLEXIBLES DE MANUFACTURA

Las nuevas tecnologías ofrecen opciones para construir y diseñar las instalaciones. Frecuentemente estas opciones son conflictivas. Por un lado, no puede enfatizarse demasiado la importancia de la flexibilidad. Especialmente en las industrias de alta tecnología, la manufactura debe adaptarse a los frecuentes avances tecnológicos. Por ejemplo, en la industria de las computadoras personales, las configuraciones que cambian de tamaño y las características del procesador central requieren que las instalaciones de manufactura constantemente se diseñen de nuevo.

Un desarrollo que ha llamado la atención de manera considerable en años recientes es el *sistema flexible de manufactura* (FMS, *flexible manufacturing system*). Un FMS es un grupo de máquinas de control numérico conectadas mediante un sistema de flujo de materiales controlado por computadora. Por lo general, las máquinas se usen para las operaciones de corte, formado y ensamblado de metales, y proporcionan el mayor beneficio cuando se requiere una gran variedad de tipos de parte. Un FMS totalmente desarrollado puede ser muy caro, llegando a requerir una inversión de capital mayor a los 10 millones de dólares. Debido al alto costo y al largo periodo de amortización, las compañías están optando por versiones reducidas, llamadas células flexibles de maquinado. Sin embargo, muchas firmas piensan que la inversión de capital se justifica. Por ejemplo, la planta Citroen en Meudon cerca de París usa un FMS para producir prototipos de componente en tamaños de lote de 1 a 50. El sistema está diseñado para manejar una amplia variedad de tipos de parte (Hartley, 1984).

¿Qué ventajas tienen estos sistemas sobre una distribución convencional de máquinas dedicadas? De acuerdo con Hartley (1984), proporcionan la oportunidad de cortar drásticamente los costos ocultos de manufactura. Estos incluyen los costos de inventario de trabajo en proceso y los costos generales asociados con la mano de obra indirecta. Permiten que las compañías cambien rápidamente las herramientas y el diseño de producto con una inversión mínima extra en equipo nuevo y personal. Sin embargo, ciertamente requieren una apreciable inversión de capital, que sólo puede recuperarse si se logra desarrollar su potencial y ambos, equipo y personal, se usan en el ambiente correcto.

La pregunta es, ¿bajo qué circunstancias debe considerar una compañía emplear un FMS como parte de su distribución general para la manufactura? Como se muestra en la figura 10-21, un sistema flexible de manufactura es apropiado cuando el volumen de producción y la variedad de las partes producidas son moderados. Para sistemas con bajo volumen y frecuente producción sobre pedido, son apropiadas las máquinas de control numérico que operan independientemente de una computadora. Estas pueden programarse para cada aplicación individual. Para la producción de alto volumen de partes estandarizadas, convienen más las líneas de transferencia fija.

La distribución y la estructura de un FMS típico se ilustran en la figura 10-22. Con frecuencia las máquinas se controlan desde una computadora central, que también puede programarse para aplicaciones individuales. Comúnmente las partes se cargan y se descargan en una sola ubicación a lo largo de la trayectoria de manejo de material. El sistema de manejo de material

FIGURA 10-21
La posición del FMS en la jerarquía de manufactura.

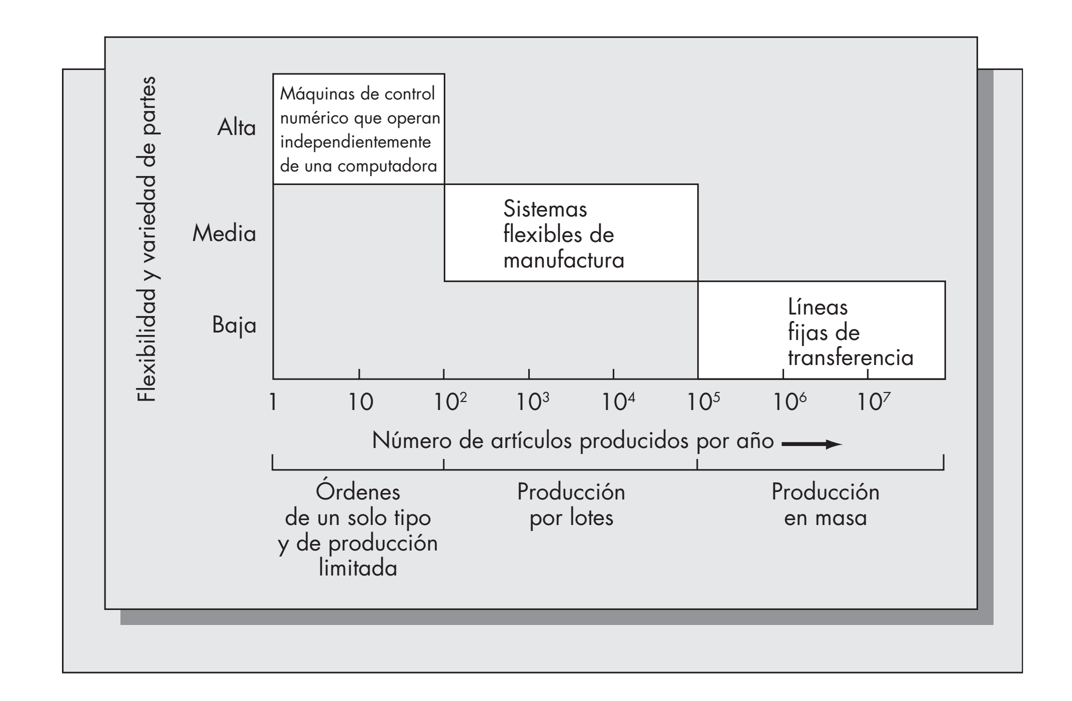
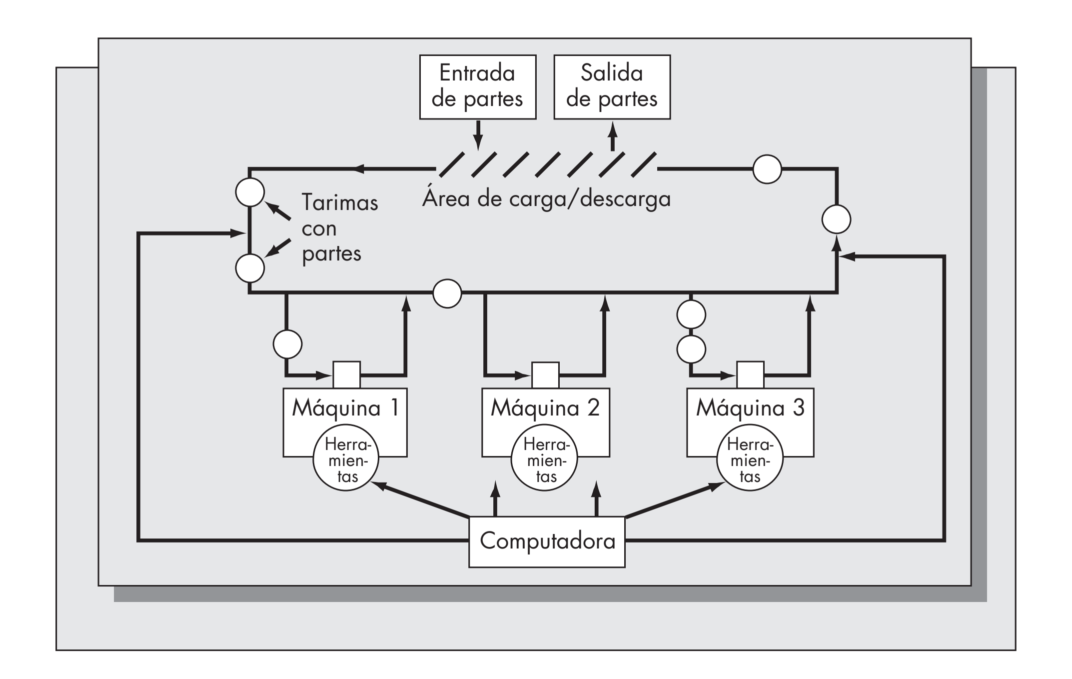


FIGURA 10-22
Un sistema flexible de manufactura típico.



consiste en tarimas, generalmente discos metálicos de dos a tres pies de diámetro. Éstas transportan inventario de trabajo en proceso y hacen cola en las máquinas para su procesamiento. Cada máquina puede contener de 5 a 100 herramientas diferentes, que se almacenan en un carrusel. Las herramientas pueden intercambiarse en cuestión de segundos. Observe que el ruteo de las partes también debe programarse en el sistema, ya que no todos los productos requieren la misma secuenciación de las operaciones de máquina. Aunque la figura 10-22 muestra una computadora centralizada, algunos sistemas usan máquinas programadas individualmente.

Ventajas de los sistemas flexibles de manufactura

Si se usan correctamente, estos sistemas pueden proporcionar ventajas sustanciales sobre diseños más rígidos. Estas incluyen

1. *Inventario reducido de trabajo en proceso*. El diseño del sistema limita el número de tarimas disponibles para mover partes a través del sistema. Así, el inventario WIP nunca sobrepasa un nivel predeterminado. En este sentido, el FMS es similar al sistema justo a tiempo, en el cual el nivel de inventario WIP es una variable de decisión cuyo valor puede escogerse anticipadamente.

2. *Utilización creciente de la máquina*. Frecuentemente las máquinas de control numérico tienen una tasa de utilización de 50 por ciento o menor. Sin embargo, un FMS eficiente puede tener una tasa de utilización tan alta como de 80 por ciento. La utilización mejorada es un resultado tanto de reducir el tiempo de cambio de los accesorios y las herramientas de la máquina, como de la capacidad para balancear mejor la carga de trabajo del sistema.

3. *Tiempo de demora reducido para la manufactura*. Sin un FMS, las partes podrían procesarse a través de varios centros diferentes de trabajo. Como resultado, podría haber un tiempo de transporte sustancial entre los centros de trabajo y un tiempo de cola sustancial en los centros de trabajo. Ya que un FMS reduce los tiempos de transporte, instalación y cambio, conduce a una reducción importante del tiempo de demora para la producción.

4. *Capacidad para manejar varias configuraciones de partes diferentes*. Como se observó anteriormente en esta sección, el FMS es más flexible que una línea fija de transferencia pero no tan flexible como una máquina de control numérico que opera independientemente de una computadora. Dependiendo de las herramientas disponibles para las máquinas, las partes pueden introducirse al sistema con poco o ningún tiempo de instalación requerido. También, el FMS puede procesar en forma simultánea las configuraciones de las partes.

5. *Costos reducidos de la mano de obra*. El número de trabajadores que se requieren para manejar un FMS puede ser 10 veces menor que el número requerido en un taller tradicional. Aun cuando las máquinas de control numérico se usen de modo que operen independientemente de una computadora, se necesita cuando menos un trabajador por máquina, y es preciso que los trabajadores transporten las partes entre las máquinas. La capacidad automatizada de manejo de material del FMS conduce a reducciones importantes en los requerimientos de la mano de obra.

Desventajas de los sistemas flexibles de manufactura

Aunque existen muchas ventajas potenciales de FMS, un factor ha demorado la aceptación de estos sistemas en Estados Unidos: el costo. La mayoría de los sistemas cuestan decenas de millones de dólares. Frecuentemente los cálculos tradicionales del valor presente neto (VPN) muestran que la inversión no se justifica. Un caso bien conocido es el de la compañía de maquinaria Yamazaki de Japón. La compañía instaló un sistema FMS de $18 millones. Como resultado, el número de máquinas se redujo de 68 a 18, el número de empleados de 215 a 12, el espacio de piso de 103000 pies cuadrados a 30000 pies cuadrados, y el tiempo promedio de procesamiento de partes de 35 días a 1.5 días (Kaplan, 1986). Estas cifras son impresionantes. Sin embargo, cuando se traducen al rendimiento de la inversión, no hay una historia color de rosa. La compañía reportó un ahorro total de $6.9 millones después de dos años. Incluyendo un ahorro de $1.5 millones por año durante los siguientes 20 años, el rendimiento total proyectado es menor al 10 por ciento por año. En la mayoría de las compañías de Estados Unidos, la tasa de rendimiento mínima aceptable es generalmente 15 por ciento o mayor, lo que convierte a la operación específica de FMS en una mala inversión sólo por las consideraciones de VPN.

Además de los costos directos del equipo y del espacio, también se incurre en varios costos indirectos. Para manejar el flujo de materiales, se necesita un sofisticado sistema de software. El software efectivo puede ser muy caro, puede requerir un diseño a la medida, con frecuencia tiene fallas de lógica y generalmente exige entrenamiento para los trabajadores. El costo de equipo tal como las alimentadoras, las bandas sin fin, y los dispositivos de transferencia pueden no

ser parte del costo inicial de compra. Otros costos indirectos incluyen la preparación del sitio, refacciones de soporte para la maquinaria e interrupciones que pueden ocurrir durante el periodo de instalación. Además, cualquier compañía que compre un FMS debe esperar el declive en la productividad que acompaña a la introducción de una nueva tecnología.

No es sorprendente que tantas compañías estadounidenses tengan problemas para justificar la inversión en FMS. Sin embargo, Kaplan (1986) arguye que los métodos tradicionales de contabilidad de costos pueden no tomar en cuenta algunas consideraciones importantes. Una es que la evaluación de las inversiones alternas basada en el flujo descontado de efectivo supone un status quo. Es decir, el análisis del valor presente neto (VPN) supone que factores tales como la participación en la plaza de mercado, el precio, los costos de mano de obra, y la posición competitiva de la compañía en el mercado permanecerán constantes. Sin embargo, los valores de algunas de estas variables cambiarán, y lo más probable es que degeneren si la compañía conserva métodos de producción obsoletos.

Además, la mayoría de las compañías prefiere invertir en varios proyectos pequeños en vez de hacer un desembolso de capital significativo para un proyecto individual. Esta filosofía es más segura a corto plazo, pero podría ser subóptima a largo plazo. Un ejemplo de una industria que no invirtió en tecnología nueva es la industria del ferrocarril. Debido a su equipo e instalaciones obsoletas, muchas compañías de esta industria no han podido ser competitivas y redituables.

El costo no es el único problema. El FMS puede experimentar lapsos de tiempo de paro de la producción para instalar nuevas operaciones o para hacer reparaciones por varias causas. Los tiempos de paro planeados podrían ser el resultado del mantenimiento programado y de los cambios programados de herramientas. Los tiempos de paro no planeados podrían resultar de fallas en las máquinas, mecánicas o eléctricas. Si una máquina individual deja de funcionar, el sistema puede continuar funcionando, pero si fallan el sistema de manejo de material o la computadora central, el FMS completo queda inutilizado.

Toma de decisiones y modelado del FMS

El modelado matemático puede ayudar con las decisiones requeridas para diseñar y administrar un FMS. Los sistemas flexibles de manufactura, al igual que todos los talleres de trabajos, deben administrarse sobre una base continua. Algunas decisiones, como sopesar la compra de un sistema, tienen consecuencias de muy largo alcance. Otras decisiones, como qué acción debe tomarse cuando se descompone una máquina o se daña una herramienta, podrían afectar solamente el flujo de materiales durante pocas horas. Esto sugiere que una manera natural de categorizar estas decisiones sería asociar el horizonte de tiempo con sus consecuencias. La tabla 10-1 resume los diferentes niveles de la toma de decisiones para FMS clasificados de esta manera.

TABLA 10-1 **Niveles de toma de decisiones en los sistemas flexibles de manufactura**

Horizonte de tiempo	Aspectos principales	Métodos de modelado
A largo plazo (meses-años)	Diseño del sistema Cambios en la mezcla de partes Modificación y/o expansión del sistema	Redes de colas Simulación Optimización
A mediano plazo (días-semanas)	Lotificación de la producción Maximización de la utilización de las máquinas Planeación de las fluctuaciones de la demanda y/o disponibilidad de los recursos	Métodos de balanceo Simulación de redes
A corto plazo (minutos-horas)	Programaciónn de las órdenes de trabajo y despacho Administración de herramientas Reacción a las fallas del sistema	Programa de operación y asignación de herramientas Algoritmo de despacho de las órdenes de trabajo Simulación

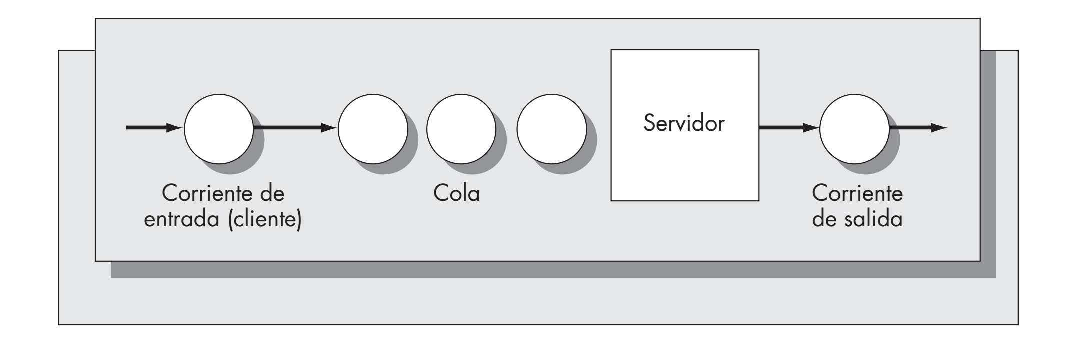


FIGURA 10-23
Una cola de un solo servidor.

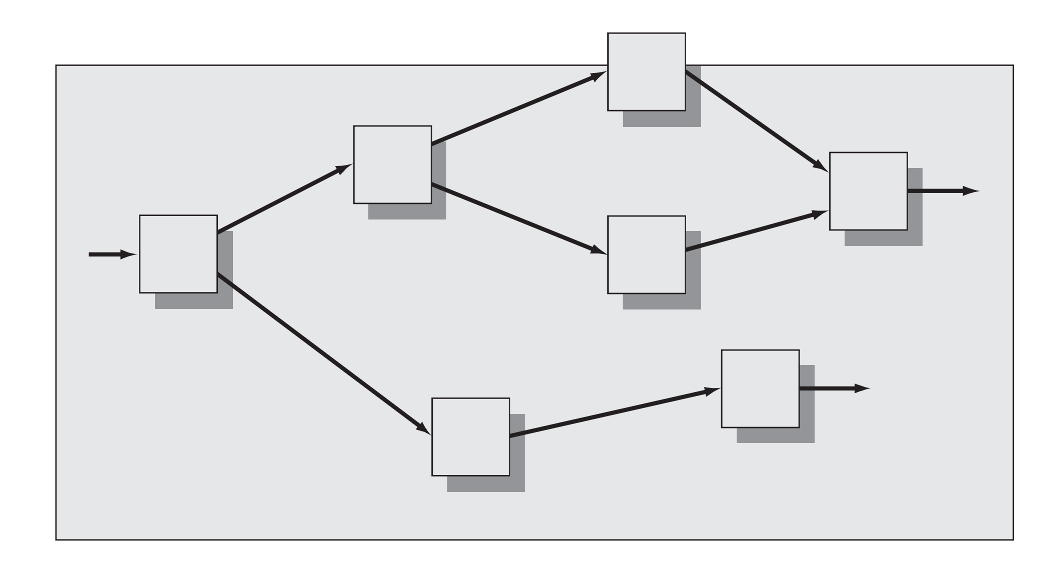


FIGURA 10-24
Una red de colas.

Modelos basados en colas

La teoría de colas es la base de muchos modelos que pueden ayudar en las decisiones a largo plazo. Una cola es simplemente una fila de espera. Un sistema de colas con un solo servidor se caracteriza por una corriente de entrada (conocida como clientes), una disciplina de servicio que especifica el orden o la secuencia de servicio para los clientes, y un mecanismo de servicio. La mayoría de los modelos de colas suponen que tanto la corriente de llegada de los clientes como los tiempos de servicio son variables aleatorias. Es la naturaleza aleatoria de las llegadas y de los servicios y sus interacciones lo que hace interesantes a los problemas de colas.[1] La figura 10-23 muestra un esquema simple de un sistema de colas con un solo servidor.

En un FMS, cada máquina corresponde a una cola separada. Los trabajos corresponden a los clientes, y la cola es la pila de trabajos que esperan su procesamiento en una máquina. Ya que un FMS es un grupo de máquinas conectadas por un sistema de manejo de materiales, la corriente de entrada de los clientes para una máquina es el resultado de la acumulación de corrientes que parten provenientes de otras máquinas. Así, un FMS es un tipo especial de red de colas. En la figura 10-24 se muestra una red de colas típica.

Los modelos de colas son muy útiles como ayuda en el diseño de sistemas. Pueden usarse para calcular varias medidas relacionadas con la utilización de las máquinas y el rendimiento del sistema. Un aspecto importante que surge durante la fase inicial de diseño es la capacidad del sistema. La capacidad del sistema es la tasa máxima de producción que puede sustentar el FMS. Depende tanto de la configuración del FMS como de las hipótesis que se propongan

[1] En el suplemento 2, que se encuentra después del capítulo 8, se muestra un resumen de la teoría básica de colas.

sobre la naturaleza de la corriente de llegada de los trabajos. El análisis de Schweitzer (1977) muestra que la tasa de salida de un sistema flexible de manufactura es simplemente la capacidad de la máquina más lenta del sistema.

Para ser más precisos, suponga que la máquina i puede procesar trabajos a una tasa μ_i. Si se conocen las probabilidades de que un trabajo que ingresa llegue a las máquinas en un orden dado, puede encontrarse e_i, el número esperado de visitas de un trabajo acabado de llegar a la máquina i, usando los métodos de la teoría de colas. Se concluye que $1/e_i$ es la tasa de llegada de los trabajos a la máquina i, de modo que μ_i/e_i es la tasa de salida de la máquina i. La máquina "más lenta", donde lento se usa en un sentido relativo, es aquella máquina cuya tasa de salida sea la menor. Así, la tasa de salida del sistema es $\text{mín}_{1\leq i\leq n}\ (\mu_i/e_i)$.

Estos resultados pueden ser muy valiosos al considerar el diseño del sistema y su utilidad potencial para la compañía. Sin embargo, en algunas circunstancias los modelos de colas tienen un sólo un valor limitado. En especial, con objeto de obtener resultados explícitos para las redes complejas de colas, con frecuencia debe suponerse que los procesos de llegada y de servicio son puramente aleatorios. Es decir, tanto el tiempo entre la llegada de los trabajos sucesivos como el tiempo requerido para que una máquina termine su tarea son variables aleatorias exponenciales.[2] Es poco probable que las distribuciones exponenciales describan con exactitud los tiempos entre llegadas y los tiempos de ejecución de los trabajos en los sistemas flexibles de manufactura. Hay circunstancias bajo las cuales las fórmulas exponenciales de colas ciertamente dan aproximaciones razonables a casos más complejos, pero debe usarse la simulación para validar las fórmulas analíticas. También se dispone de resultados analíticos aproximados para los casos no exponenciales (véanse Suri y Hildebrant, 1984, por ejemplo). Asimismo, el supuesto exponencial no se requiere para algunas configuraciones simples de red. En Buzacott y Yao (1986), puede encontrarse una revisión de los resultados analíticos para modelos de red de colas de los sistemas flexibles de manufactura.

Aun con estas limitaciones, los modelos de colas proporcionan un medio poderoso para analizar las características de desempeño y limitaciones de capacidad de los sistemas flexibles de manufactura. Finalmente, los modelos de colas en red reflejarán con exactitud la naturaleza de los FMS y proporcionarán un medio efectivo para ayudar en el diseño y la administración del sistema.

Modelos de programación matemática

Stecke (1983) considera el uso de la programación matemática para solucionar varios problemas de decisiones que surgen de la administración del FMS. Estos problemas incluyen los siguientes:

1. *Problema de la selección del tipo de parte*. A partir de un conjunto dado de tipos de parte que poseen requerimientos conocidos asociados, determine un subconjunto para el procesamiento inmediato y simultáneo.
2. *Problema del agrupamiento de las máquinas*. Divida las máquinas en grupos de modo que cada máquina de un grupo pueda ejecutar el mismo conjunto de operaciones.
3. *Problema de la razón de producción*. Determine las razones relativas para producir los tipos de parte seleccionados en el paso 1.
4. *Problema de la asignación de recursos*. Asigne el número limitado de tarimas y soportes entre los tipos seleccionados de partes.
5. *Problema de la carga*. Asigne las operaciones y las herramientas requeridas de los tipos seleccionados de partes entre los grupos de máquinas sujetas a restricciones de capacidad tecnológica del FMS.

[2] Las propiedades de la distribución exponencial se discuten en detalle en el capítulo 2 y en el suplemento 2 sobre colas.

No presentaremos los detalles de las formulaciones de programación entera mixta. Nuestro propósito es simplemente observar que la programación matemática es otro medio que ayuda a las decisiones a largo plazo y a mediano plazo que afectan al FMS. Stecke (1983) muestra cómo se aplicarían estos modelos usando datos reales provenientes de un FMS existente.

Métodos de simulación

La simulación por computadora ha sido desde hace mucho tiempo una herramienta popular y útil para analizar sistemas complejos de manufactura que no se analizan fácilmente con medios más convencionales. Una simulación por computadora es un programa diseñado para modelar las características importantes de un sistema real. La simulación Monte Carlo implica que existen elementos de aleatoriedad en el problema. En el FMS la aleatoriedad surge de la variación en la corriente de entrada, la variación en la mezcla de partes, y posibles fallas en las máquinas, en las herramientas, en las computadoras o en el sistema de manejo de materiales. Como el desarrollo de modelos de simulación es costoso y puede requerir mucho tiempo de cómputo, se prefieren los modelos analíticos exactos. Sin embargo, con frecuencia los supuestos requeridos por los modelos de colas o la programación matemática son demasiado restrictivos. En estos casos la simulación puede ser la única alternativa.

Los programas de simulación pueden escribirse en un lenguaje fuente tal como Fortran o BASIC. Sin embargo, existen lenguajes de simulación de uso especial que son más populares para las aplicaciones de manufacturas. Ejemplos típicos de estos lenguajes son GPSS y SLAM. Este último resulta especialmente adecuado para simular los tipos de redes que surgen en las aplicaciones de manufactura. Las simulaciones son útiles para responder a preguntas del tipo "qué pasa si". Pritsker (1986) describe el uso de la simulación específicamente para el diseño y la evaluación de un FMS usado para controlar las operaciones de maquinado en cierto tipo de fundición de metal. Los objetivos del estudio de simulación eran evaluar el balance y la productividad del sistema, examinar la necesidad de equipo adicional, determinar los recursos, si se requieren algunos, cuáles deben eliminarse, y encontrar el número de maquinas requeridas tanto dedicadas como flexibles. La simulación reveló que parte del equipo existente podría eliminarse alcanzándose de todos modos la meta deseada de producción.

Un paquete de simulación diseñado específicamente para la modelación de fábricas es XCELL (véanse Conway y otros, 1987). El piso de la fábrica se representa como una malla uniforme de celdas y cada "elemento" de la fábrica ocupa una celda. Un elemento puede ser 1) un centro de trabajo, 2) un área de recepción, 3) un área de embarque, 4) un tope, o 5) una instalación de mantenimiento. XCELL emplea representaciones gráficas simbólicas de los elementos. Aunque XCELL todavía no logra la aceptación de GPSS y SLAM, es fácil de usar y la flexibilidad para la modelación del flujo de fábrica lo hace una opción atractiva.

Un nuevo desarrollo en la simulación de FMS ha sido la aplicación de técnicas desarrolladas para evaluar el rendimiento de las redes de computadoras y de telecomunicaciones. Bruno y Biglia (1985) exploran la aplicación de las redes estocásticas de Petri a algunos problemas de administración en FMS. Las redes de Petri son representaciones en red de sistemas con características especiales tales como distintivos que indican cuándo se bloquean ciertas partes del sistema. En un FMS, puede ocurrir un bloqueo si una máquina se descompone o si el número de partes que esperan ser procesadas en una máquina es igual al tamaño máximo de la cola. La aplicación de la teoría de redes de Petri a los problemas de diseño y administración de FMS es un área abierta a la investigación. Todavía queda por ver si esta teoría va a proporcionar algún resultado útil para la administración o el diseño de FMS que no pueda lograrse mediante métodos de simulación más convencionales.

El futuro del FMS

Existe un potencial importante en la aplicación del FMS tanto en Estados Unidos como en el exterior. Se estima que aproximadamente la mitad de los gastos anuales en manufactura radica en trabajar los metales, y dos tercios de dicha industria se abocan al corte de los mismos (Stecke, 1983). Además, aproximadamente tres cuartos del volumen en dólares de las partes de

metal trabajado se manufacturan en lotes de menos de 50 partes (Cook, 1975). Evidentemente, hay un enorme mercado para los FMS en Estados Unidos. También existe un potencial similar en otros países industrializados. De hecho, la tasa de crecimiento de los FMS instalados es más alta en Japón que en cualquier otro país del mundo (Jaikumar, 1984).

La tasa de crecimiento de los FMS en Estados Unidos ha sido saludable si no es que espectacular. En 1981 hubo más o menos 18 sistemas operativos en Estados Unidos (Jaikumar, 1984), y a finales de 1986 este número había crecido aproximadamente hasta 50. Krouse (1986) estimó que debió haber habido 284 sistemas instalados en Estados Unidos para 1990. Como se observó anteriormente, la principal preocupación entre las compañías que consideran comprar un FMS es el precio. El costo promedio de estos sistemas se estimó en $12.5 millones en 1986 y es probable que sea más alto en la actualidad. También, la instalación generalmente requiere de 18 a 24 meses.

Los costos del hardware y de la conversión desalientan a muchos usuarios potenciales de FMS. Además, a muchas fábricas en Estados Unidos todavía les falta invertir en máquinas con herramientas de control numérico, para no hablar de un FMS totalmente desarrollado. En respuesta, los vendedores están ofreciendo versiones reducidas del FMS conocidas como células flexibles de manufactura. Las células flexibles de manufactura son FMS pequeños que ofrecen menos máquinas, un sistema de manejo de material menos extenso, y menos herramientas por máquina. Los vendedores esperan ventas saludables de estas células, que conducirían a expandir la base de clientes y a un mercado más grande para sistemas escalados.

El costo no es el único factor que retrasa la aceptación. El desempeño de muchos FMS existentes ha sido decepcionante. Zygmont (1986) reportó que algunos sistemas instalados no cumplieron las expectativas de las compañías que los adquirieron. En un caso, se necesitaron tres años para corregir los errores de cómputo del sistema que compró Electro-Optical and Data Systems de Hughes Aircraft, y al final el sistema era todavía menos flexible que lo esperado originalmente. Uno de los problemas encontrados fue que el nivel de precisión requerido era más alto que lo que era capaz de entregar el sistema. En forma similar, Deere and Cose decepcionó con su sistema de $20 millones, y quedó especialmente desilusionado ante la falta de flexibilidad que presentaba el software para manejar secuencias de partes complejas y variables.

Jaikumar (1986) afirma que el problema con los FMS en Estados Unidos no son los sistemas en sí mismos sino la manera en que se usan. Al comparar los FMS de Estados Unidos con los de Japón, encontró algunas diferencias notables. Estas diferencias se listan en la tabla 10-2.

Resulta significativo el hecho de que hay casi 10 veces más tipos de partes producidas en cada sistema japonés que en cada sistema estadounidense. Eso, junto con la enorme diferencia en el volumen anual por parte, indica que las compañías estadounidenses están usando inapropiadamente los FMS. Es decir, los usan como si fueran simplemente otro conjunto de máquinas para la producción estandarizada de alto volumen en vez de producir la mezcla de diferentes partes para lA que fueron diseñados. Otra diferencia importante es que muchos sistemas japoneses

TABLA 10-2 **Comparación de los FMS en Estados Unidos y en Japón**

	Estados Unidos	Japón
Veces de desarrollo de sistemas (años)	25 a 3	1 a 1.25
Tipos de partes producidas por sistema	10	93
Volumen anual por parte	1.727	258
Número de partes nuevas introducidas por año	1	22
Número de sistemas con operaciones sin supervisión	0	18
Tiempo promedio dedicado al corte de metal por día	83	202

corren día y noche y frecuentemente sin supervisión. Jaikumar (1986) afirma que ésta es una consecuencia del diseño mejorado y de la confiabilidad de los sistemas japoneses.

Conforme las compañías instalen FMS adicionales, los vendedores comprenderán mejor su potencial y sus limitaciones. Los avances tanto en el hardware como en el software deben hacer que los sistemas futuros tengan mayor capacidad de manejo de los problemas estudiados. Según mejoren los sistemas y la comprensión de las circunstancias bajo las cuales pueden ser muy efectivos, veremos a muchas más compañías instalando FMS en la siguiente década.

Problemas para la sección 10.7

20. En cada caso listado, diga si cada factor es una ventaja o una desventaja de los FMS. Discuta las razones que lo llevaron a tal elección.

    a) Costo.
    b) Capacidad para manejar requerimientos de partes diferentes.
    c) Avances en la tecnología de manufactura.
    d) Confiabilidad.

21. Para cada una de las situaciones de caso descritas, diga cuál de los tres tipos de sistemas de manufactura sería el más apropiado: (1) máquinas de control numérico que operan independientemente de una computadora, (2) FMS, (3) línea de transferencia fija, (4) otro sistema. Explique su selección en cada caso.

    a) Un taller local de máquinas que acepta órdenes a la medida de diferentes clientes de la industria electrónica.
    b) Un proveedor internacional de contenedores para metal tamaño estándar.
    c) La división de trabajado de metales de un fabricante grande de aviones, que debe satisfacer las necesidades de varias divisiones en la compañía.

22. ¿Por qué razón un VPN (valor presente neto) tradicional podría no ser un medio apropiado para evaluar la conveniencia de comprar un FMS?
23. ¿Para cuáles decisiones relacionadas con el diseño o el control de un FMS pueden ser útiles los siguientes métodos? También, describa brevemente cómo se usaría cada método.

    a) Teoría de colas
    b) Simulación
    c) Programación matemática

10.8 UBICACIÓN DE NUEVAS INSTALACIONES

En el capítulo 1 se discutieron varias consideraciones cualitativas en torno a decidir la mejor ubicación de una nueva planta o de otra instalación. En el resto de este capítulo se considerarán las técnicas cuantitativas que puedan ayudar a decidir la ubicación. Existen muchas circunstancias para las cuales el objetivo puede cuantificarse fácilmente, y es en estos casos cuando los métodos de solución analítica son muy útiles.

En cierto sentido, el problema de la distribución de planta es un caso especial del problema de ubicación. Sin embargo, los problemas considerados en el resto de este capítulo difieren de los problemas de distribución por lo siguiente: supondremos que hay una o más instalaciones existentes que ya están en su lugar y deseamos encontrar la ubicación óptima de una instalación nueva. Nos concentraremos en el problema de la ubicación de una nueva instalación individual, aunque existen versiones del problema en las cuales deben determinarse las ubicaciones de instalaciones múltiples.

Aplicación instantánea

KRAFT FOODS UTILIZA LA OPTIMIZACIÓN Y LA SIMULACIÓN PARA DETERMINAR LA MEJOR DISTRIBUCIÓN

A mediados de la década de 1990 Kraft decidió renovar una de sus principales instalaciones de manufactura, ubicada en el Medio Oeste. Nuevas líneas de producción con mayor capacidad debían remplazar a las antiguas. Las nuevas líneas tendrían mayor producción y más capacidad de mezcla, y podrían operar a diferentes velocidades. Una de las preocupaciones en el diseño de la nueva instalación era conservar los niveles presentes de capacidad al tiempo que también se permitiera un futuro crecimiento. La gerencia necesitaba decidir rápidamente la configuración de la línea porque se requería un tiempo de entrega de seis meses para dichas líneas.

La gerencia de Kraft pidió al grupo de investigación de operaciones de la compañía que desarrollara un modelo matemático detallado del sistema. Un enfoque improvisado era simplemente inaceptable: una decisión equivocada sería costosa. Demasiadas líneas conducirían a un desperdicio de espacio y de dinero, y demasiado pocas a un cuello de botella en la capacidad. Una dificultad adicional era que ninguna de las plantas de Kraft había usado antes estas líneas, por lo que no había experiencia previa en la cual apoyarse.

Se pidió a los miembros del grupo de investigación de operaciones de la compañía que considerara tanto el problema de optimizar el número de las nuevas líneas como el flujo de material dentro de la planta. Para encarar el problema de optimización, el grupo desarrolló una formulación de programación entera del problema. El objetivo era asegurar que la terminación de cualquier calendario que posiblemente encontrara Kraft, no sobrepasara un turno de ocho horas. (Consulte el capítulo 8 para un análisis sobre la terminación.) Se determinó que seis líneas fácilmente se encargarían de los requerimientos existentes de la capacidad y proporcionarían un margen sustancial para un futuro crecimiento. El modelo se resolvió usando el paquete de optimización AMPL en una computadora personal.

Una vez que se determinó el número óptimo de líneas nuevas, el siguiente paso era desarrollar una simulación detallada del piso de fábrica para mostrar gráficamente cómo trabajaría en la práctica el nuevo diseño. Para dicha tarea, el grupo decidió usar el paquete de software Taylor II de F&H Simulation, un paquete de simulación basado en gráficos especialmente adecuado para considerar los efectos de diferentes programas de trabajo en una distribución dada. El modelo de simulación constaba de 139 elementos incluyendo líneas, máquinas, amortiguadores de trabajo en proceso, áreas de carga para la materia prima, y áreas de embarque. Al simular el proceso con los calendarios más complejos, el grupo demostró que la nueva distribución tendría suficiente capacidad para manejar las demandas más severas de las que la planta posiblemente afrontaría. Además, la simulación ofreció al personal de la fábrica un medio para ver cómo cambiaría el flujo de materiales antes de renovar la planta e implementar la nueva distribución.

Fuente: Basado en Sengupta y Combes (1995).

Los ejemplos del tipo de problemas de ubicación de una instalación que pueden resolverse mediante los métodos estudiados en este capítulo incluyen

1. Ubicar un nuevo almacén para una compañía con una red existente de centros de producción y distribución.
2. Determinar la mejor ubicación de una nueva máquina en un taller de trabajos.
3. Situar el centro de cómputo en una nueva universidad.
4. Encontrar la mejor ubicación de un nuevo hospital en un área metropolitana.
5. Encontrar la ubicación más adecuada de una planta de generación de energía diseñada para servir a una región geográfica.
6. Determinar dónde colocar un cajero automático en un vecindario.
7. Ubicar una nueva estación de policía en una comunidad.

Indudablemente, el lector puede pensar en muchos otros ejemplos de problemas de ubicación. Los métodos analíticos suponen que el objetivo es situar la nueva instalación para minimizar alguna función de la distancia de la nueva ubicación a partir de ubicaciones existentes.

Por ejemplo, cuando se ubica un almacén nuevo, un objetivo apropiado sería minimizar la distancia total recorrida al almacén desde las instalaciones de producción, y desde el almacén hasta los subdistribuidores minoristas. Un hospital se ubicaría de modo que fuera de fácil acceso a la mayor proporción de la población en el área de servicio. Una máquina se colocaría de modo que se minimizara la suma ponderada de viajes de manejo de material desde y hacia la máquina. Es evidente que seleccionar la ubicación de una instalación para minimizar alguna función de la distancia que separa a la instalación nueva de las instalaciones existentes resulta apropiado para muchos problemas reales de localización.

Medidas de distancia

Son muy comunes dos medidas de distancia: *la distancia euclidiana* y *la distancia rectilínea*. La distancia euclidiana también se conoce como distancia en línea recta. La distancia euclidiana que separa dos puntos es simplemente la longitud de la línea recta que une dichos puntos. Suponga que una instalación existente se ubica en el punto (a, b) y sea (x, y) la ubicación de la nueva instalación. Entonces la distancia euclidiana entre (a, b) y (x, y) es

$$\sqrt{(x - a)^2 + (y - b)^2}.$$

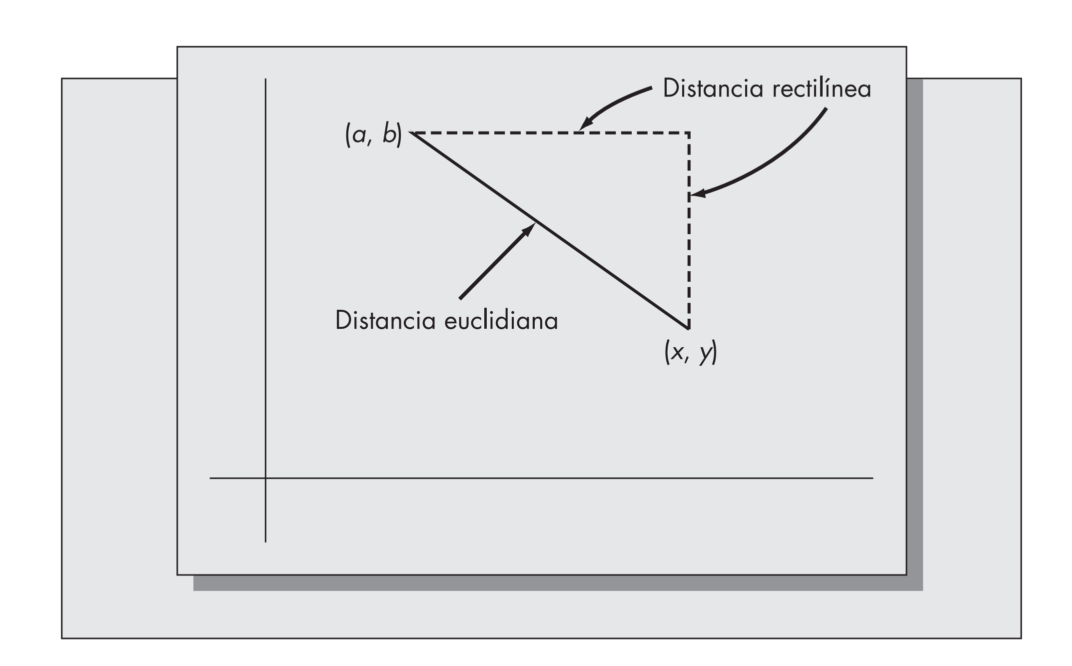
La distancia rectilínea (también conocida como la distancia metropolitana, reconociendo el hecho de que las calles generalmente siguen un patrón reticular) está dada por la fórmula

$$|x - a| + |y - b|.$$

La figura 10-25 ilustra la diferencia entre estas dos medidas de distancia.

La distancia rectilínea es apropiada para muchos problemas de ubicación. Las distancias en las áreas metropolitanas tienden a aproximarse más a las distancias rectilíneas que a las distancias euclidianas aun cuando el patrón de las calles no sea una retícula perfecta. En muchos ambientes de manufactura, el material se transporta a través de corredores arreglados según cierto patrón regular. El hecho de que la distancia rectilínea sea más común que la distancia euclidiana es una coincidencia afortunada, porque el problema de la distancia rectilínea es más fácil de resolver.

FIGURA 10-25
Distancias euclidiana y rectilínea.



Problemas para la sección 10.8

24. Considere el problema de ubicar un nuevo hospital en un área metropolitana. Liste los factores que pueden cuantificarse y los que no. Comente sobre la utilidad de los métodos cuantitativos en la toma de decisiones.
25. Un sistema coordenado se superpone sobre un mapa. Tres instalaciones existentes se ubican en (5, 15), (10, 20), y (6, 9). Calcule las distancias rectilínea y euclidiana que separan cada instalación de una nueva instalación ubicada en $(x, y) = (8, 8)$.
26. Para la situación descrita en el problema 25, suponga que existen solamente tres ubicaciones factibles para tal distribución: (8, 16), (6, 15), y (4, 18).

    a) ¿Cuál es la ubicación óptima si el objetivo es minimizar la distancia rectilínea total hacia las tres instalaciones existentes?

    b) ¿Cuál es la ubicación óptima si el objetivo es minimizar la distancia euclidiana total hacia las tres instalaciones existentes?

27. Para cada uno de los siete ejemplos de problemas de ubicación listados en esta sección, indique cuál medida de distancia, euclidiana o rectilínea, sería la más apropiada. (Discuta cómo, en algunos casos, uno u otro objetivo serían apropiados para el mismo problema, según el criterio de optimización que se use.)

10.9 EL PROBLEMA DE UBICACIÓN DE UNA SOLA INSTALACIÓN CON DISTANCIA RECTILÍNEA

En esta sección presentaremos una solución al problema general de ubicar una nueva instalación entre n instalaciones existentes. El objetivo es ubicar la nueva instalación de modo que se minimice la suma ponderada de las distancias rectilíneas desde la nueva instalación hacia las instalaciones existentes. Suponga que las instalaciones existentes se ubican en los puntos (a_1, b_1), (a_2, b_2), . . . , (a_n, b_n). Entonces la meta es encontrar valores de x e y para minimizar

$$f(x, y) = \sum_{i=1}^{n} w_i(|x - a_i| + |y - b_i|).$$

Los pesos se incluyen para permitir diferentes tasas de tránsito entre la nueva instalación y las instalaciones existentes. Una propiedad simplificatoria del problema es que los valores óptimos de x e y pueden determinarse por separado, como

$$f(x, y) = g_1(x) + g_2(y).$$

donde

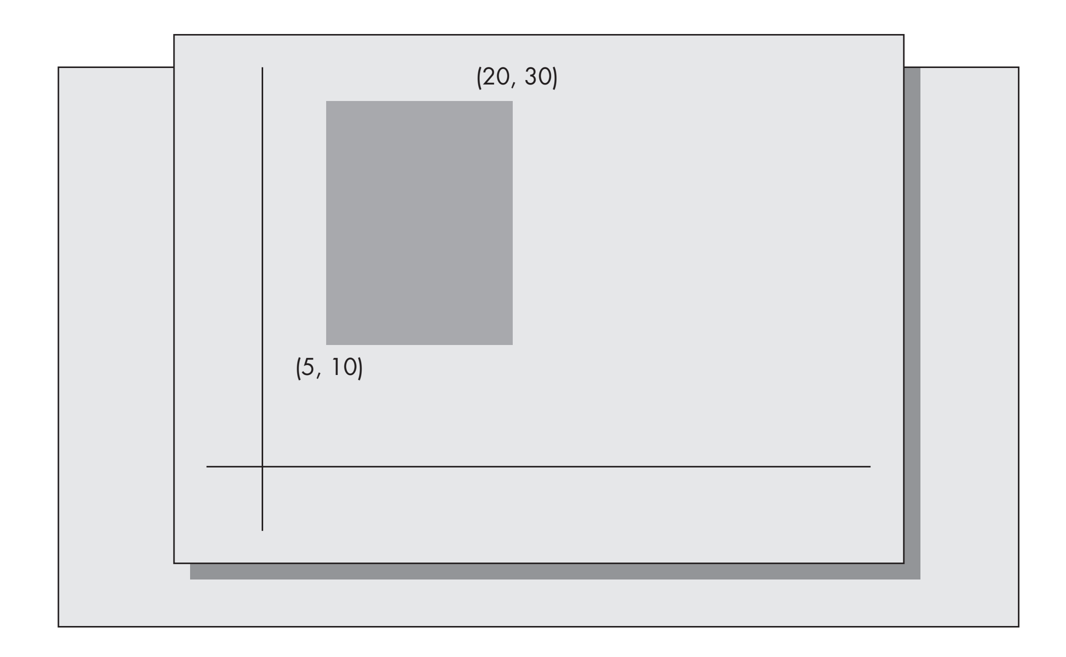
$$g_1(x) = \sum_{i=1}^{n} w_i|x - a_i|$$

y

$$g_2(y) = \sum_{i=1}^{n} w_i|y - b_i|.$$

Según veremos, siempre existe una solución óptima con x igual a algún valor de a_i e y igual a algún valor de b_i. (También puede haber otras soluciones óptimas.)

FIGURA 10-26
Ubicaciones óptimas de la nueva instalación para una medida de distancia rectilínea.



Sin embargo, antes de presentar el algoritmo de la solución general para encontrar la ubicación óptima de la nueva instalación, consideremos algunos ejemplos simples con objeto de proporcionar al lector alguna intuición. Considere primero el caso en el cual hay exactamente dos instalaciones existentes, ubicadas en (5, 10) y (20, 30) como se ilustra en la figura 10-26. Suponga que el peso aplicado a cada una de estas instalaciones es 1. Si x adopta cualquier valor entre 5 y 20, el valor de $g_1(x)$ es igual a 15. (Por ejemplo, si $x = 13$, entonces $g_1(x) = |5 - 13| + |13 - 20| = 8 + 7 = 15$.) En forma similar, si y adopta cualquier valor entre 10 y 30, entonces $g_2(y)$= 20. Cualquier valor de x fuera del intervalo cerrado [5, 20] y cualquier valor de y fuera del intervalo cerrado [10, 30] conducen a valores mayores de $g_1(x)$ y $g_2(y)$. Así, la solución óptima es (x, y) con $5 \leq x \leq 20$ y $10 \leq y \leq 30$. Todas las ubicaciones en la región sombreada que se ilustra en la figura 10-26 son óptimas.

Como en el ejemplo, siempre habrá una ubicación óptima de la nueva instalación con coordenadas que provienen del conjunto de coordenadas de las instalaciones existentes. Suponga que las instalaciones existentes tienen las ubicaciones (3, 3), (6, 9), (12, 8), y (12, 10). Nuevamente suponga que el peso aplicado a estas ubicaciones es 1. La jerarquización de las ubicaciones x en orden creciente da 3, 6, 12, 12. Un valor *mediano* es tal que la mitad de los valores de x están situados arriba de éste y la mitad de los valores de x están situados debajo de éste. Cualquier valor de x entre 6 y 12 es una posición de la mediana y es óptimo para este problema. El valor óptimo de $g_1(x)$ es 15. (El lector debe experimentar con un número de valores diferentes de x entre 6 y 12 para cerciorarse de que éste sea el caso.) La jerarquización de los valores de y en orden ascendente da 3, 8, 9, 10. El valor mediano de y está entre 8 y 9, y el valor óptimo de $g_2(y) = 8$.

La solución óptima es ubicar (x, y) en la mediana de las instalaciones existentes. Este resultado se amplía al caso en el cual hay pesos diferentes de 1. Suponga en el ejemplo 10.4 que nos dieran las ubicaciones de cuatro máquinas en un taller de trabajos. La meta es encontrar la ubicación de una quinta máquina para minimizar la distancia total recorrida para transportar material entre la nueva máquina y las existentes. Asuma que en promedio hay respectivamente

2, 4, 3, y 1 viajes de manejo de material por hora desde las máquinas existentes a la nueva máquina. Resumiendo la información dada,

Ubicación de las máquinas existentes	Peso
(3, 3)	2
(6, 9)	4
(12, 8)	3
(12, 10)	1

Este problema es equivalente a aquel en el cual hay dos máquinas en la ubicación (3, 3), cuatro máquinas en la ubicación (6, 9), tres máquinas en la ubicación (12, 8), y una máquina en la ubicación (12, 10), con pesos iguales a 1. Así, las ubicaciones x en orden creciente son 3, 3, 6, 6, 6, 6, 12, 12, 12, 12. La ubicación mediana es $x = 6$. Las ubicaciones y en orden creciente son 3, 3, 8, 8, 8, 9, 9, 9, 9, 10. La ubicación mediana es cualquier valor de y en el intervalo [8, 9]. El lector debe verificar que el valor de la función objetivo para la solución óptima es $30 + 16 = 46$.

Si los pesos son grandes, este enfoque es inconveniente. Un método más rápido para encontrar la ubicación óptima de la nueva instalación es calcular los pesos acumulados y determinar la ubicación o las ubicaciones correspondientes a la mitad del peso acumulado. El procedimiento se ilustra mejor mediante un ejemplo.

Ejemplo *10.5*

La Universidad del Lejano Oeste ha comprado equipo que permite a los profesores filmar videos de las conferencias. El equipo será usado por los profesores de seis escuelas en el campus: negocios, educación, ingeniería, humanidades, leyes, y ciencias. Las ubicaciones de los edificios en el campus se ilustran

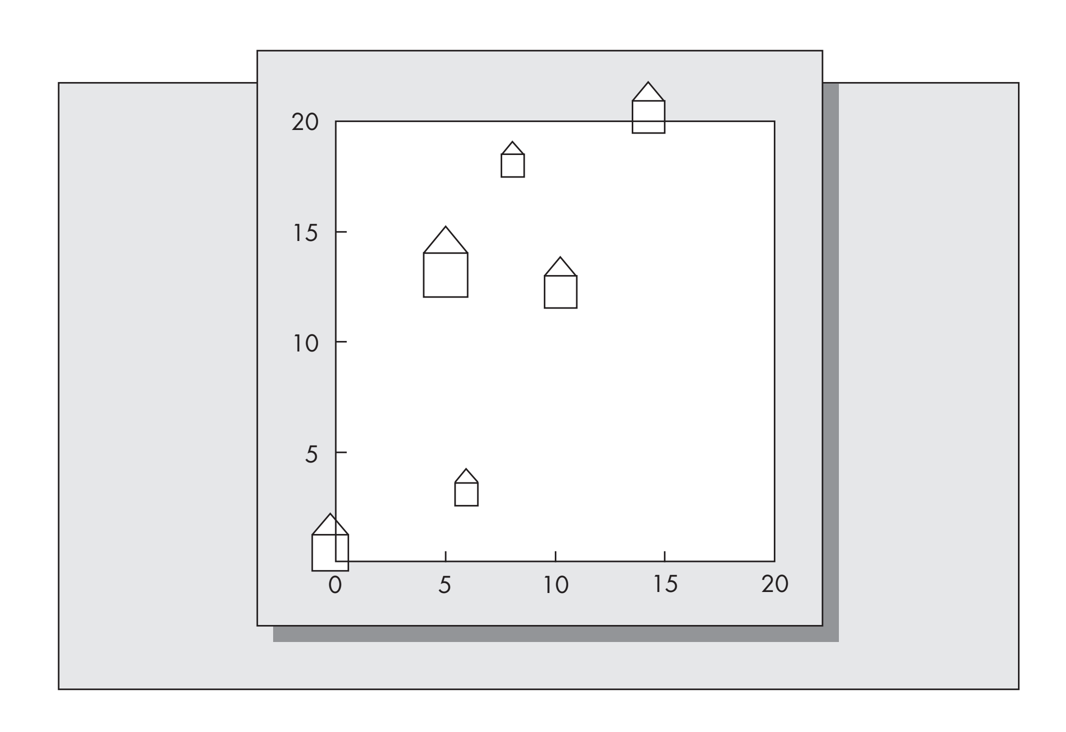


FIGURA 10-27
Ubicación de seis edificios en el campus (véase el ejemplo 10.5).

en la figura 10-27. Las coordenadas de las ubicaciones y el número de profesores que se espera usen el equipo son los siguientes:

Escuela	Ubicación en el campus	Número de profesores
Negocios	(5, 13)	31
Educación	(8, 18)	28
Ingeniaría	(0, 0)	19
Humanidades	(6, 3)	53
Leyes	(14, 20)	32
Ciencias	(10, 12)	41

El campus posee grandes áreas con césped que separan a los edificios, y los andadores son principalmente este-oeste o norte-sur, de modo que las distancias entre los edificios son rectilíneas. Al planificador de la universidad le gustaría ubicar la nueva instalación de modo que se minimice el tiempo total de recorrido de todos los profesores que piensen usarla.

Encontraremos los valores óptimos de las coordenadas x e y por separado. Considere el valor óptimo de la coordenada x. Primero jerarquizamos las coordenadas x en valor creciente y acumulamos los pesos.

Escuela	Coordenada *x*	Peso	Peso acumulado
Ingeniería	0	19	19
Negocios	5	31	50
Humanidades	6	53	103
Educación	8	28	131
Ciencias	10	41	172
Leyes	14	32	204

El valor óptimo de la coordenada x se encuentra dividiendo el peso acumulado total entre 2 e identificando la primera ubicación para la cual el peso acumulado sobrepase a este valor. En el ejemplo ésta es la primera vez que el peso acumulado sobrepasa a $204/2 = 102$. Esto ocurre en $x = 6$, cuando el peso acumulado es 103. Así, la x óptima es igual a 6.

Usamos el mismo procedimiento para encontrar el valor óptimo de la coordenada y. Las jerarquizaciones se dan en la siguiente tabla.

Escuela	Coordenada *y*	Peso	Peso acumulado
Ingeniería	0	19	19
Negocios	3	53	72
Humanidades	12	41	113
Educación	13	31	144
Ciencias	18	28	172
Leyes	20	32	204

En este caso el peso acumulado sobrepasa primero a 102 cuando $y = 2$. Así, $y = 12$ es óptimo. La ubicación óptima de la nueva instalación es (6, 12). Esta solución es única. Se presentarán óptimos múltiples cuando un valor del peso acumulado sea exactamente igual a la mitad del peso acumulado total. Por ejemplo, suponga que el peso de Ciencias fuera 19 en lugar de 41. Entonces el peso acumulado de Ciencias sería 91, exactamente la mitad del peso total de 182, y todos los valores de y en el intervalo cerrado [12, 13] serían óptimos.

Curvas de nivel

El ejemplo 10.5 sugiere una pregunta interesante. Suponga que la ubicación (6, 12) es inverosímil. ¿Cómo puede medir la universidad la penalidad en el costo si la nueva instalación se ubica en otro lado? Las curvas de nivel, o curvas de isocosto, pueden ayudar a determinar la penalidad de las soluciones no óptimas. Una curva de nivel es una curva de costo constante: la ubicación de una nueva instalación en cualquier punto a lo largo de una curva de nivel conduce exactamente al mismo costo.

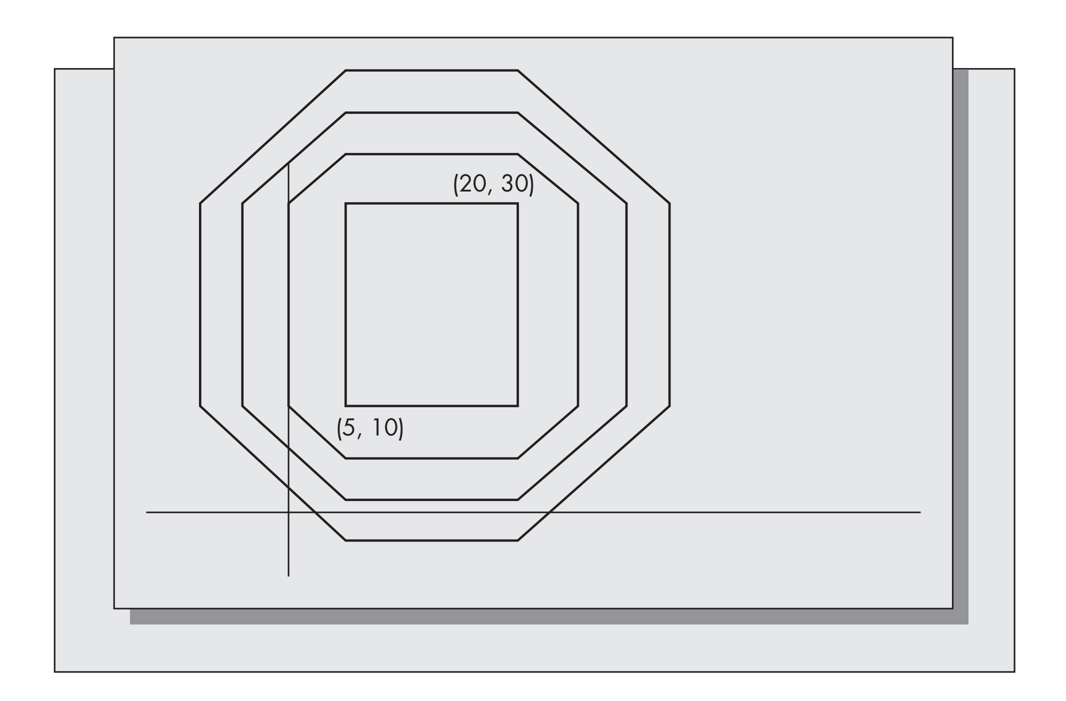


FIGURA 10-28
Curvas de nivel para el problema de dos instalaciones ilustrado en la figura 10-26.

En la figura 10-28 hemos ilustrado las curvas de nivel para el ejemplo sencillo de la figura 10-26, en el cual hay solamente dos instalaciones existentes y los pesos son iguales. El lector debe convencerse de que la distancia rectilínea total desde cualquier punto a lo largo de una curva de nivel hasta los dos puntos (5, 10) y (20, 30) es la misma. La determinación de las curvas de nivel incluye calcular la pendiente apropiada para cada una de las regiones obtenidas al trazar líneas verticales y horizontales que pasen por cada uno de los puntos (a_i, b_i). El procedimiento para determinar las curvas de nivel se esboza en el apéndice 10-B al final de este capítulo.

En la figura 10-29 hemos ilustrado las curvas de nivel para el problema del ejemplo 10.5 en el cual la universidad debe ubicar un centro audiovisual. Si la ubicación óptima (6, 12) es inverosímil, la administración de la universidad podría usar este mapa para ver las penalidades asociadas con sitios alternos.

Problemas Minimax

Hasta ahora hemos supuesto que la nueva instalación debe disponerse para reducir la suma de las distancias ponderadas a todas las instalaciones existentes. Sin embargo hay circunstancias para las cuales este objetivo es inapropiado. Piense en el siguiente ejemplo. La ciudad está considerando las ubicaciones para una instalación paramédica. Los paramédicos deben tener capacidad para responder a las llamadas de emergencia desde cualquier parte de la ciudad. Ciertas condiciones, tales como un grave ataque cardiaco, deben tratarse rápidamente para que el paciente tenga alguna posibilidad de supervivencia. Así, la instalación debe ubicarse de modo que *todas* las ubicaciones en la ciudad puedan alcanzarse en un tiempo dado.

En este caso el objetivo sería determinar dónde ubicar la nueva instalación para minimizar la distancia máxima hasta las instalaciones existentes en lugar de la distancia total. Sea $f(x, y)$ la distancia máxima desde la nueva instalación hasta las instalaciones existentes. Entonces

$$f(x, y) = \max_{1 \le i \le n} (|x - a_i| + |y - b_i|).$$

El objetivo es encontrar (x^*, y^*) que satisfaga

$$f(x^*, y^*) = \min_{x,y} f(x, y).$$

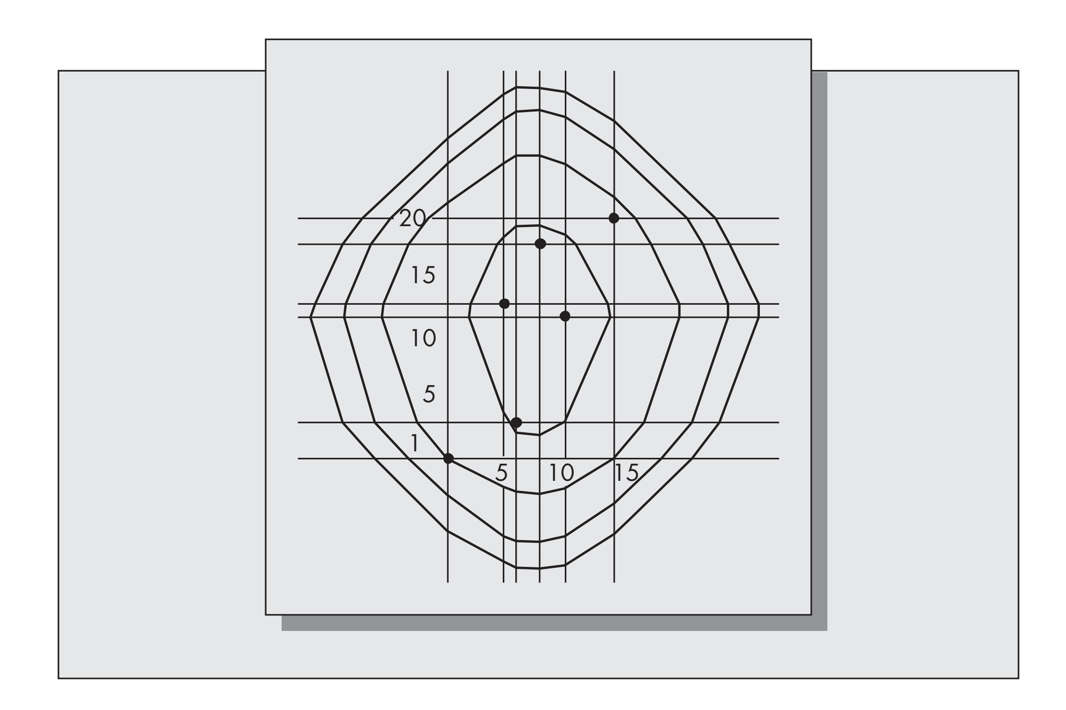


FIGURA 10-29
Curvas de nivel para la ubicación de la universidad en el ejemplo 10.5.

El procedimiento para encontrar la ubicación minimax óptima es directo. Puede usarse la programación lineal para mostrar que el procedimiento que esbozaremos es óptimo. (No presentaremos los detalles aquí. El lector interesado puede remitirse a Francis y White, 1974.) Defina los números c_1, c_2, c_3, c_4, y c_5:

$$c_1 = \min_{1\le i\le n}(a_i + b_i),$$
$$c_2 = \max_{1\le i\le n}(a_i + b_i),$$
$$c_3 = \min_{1\le i\le n}(-a_i + b_i),$$
$$c_4 = \max_{1\le i\le n}(-a_i + b_i),$$
$$c_5 = \max(c_2 - c_1, c_4 - c_3).$$

Sea

$$x_1 = (c_1 - c_3)/2,$$
$$y_1 = (c_1 + c_3 + c_5)/2,$$

y

$$x_2 = (c_2 - c_4)/2,$$
$$y_2 = (c_2 + c_4 - c_5)/2.$$

Entonces todos los puntos que se encuentran a lo largo de la línea que conecta a (x_1, y_1) y (x_2, y_2) son óptimos. Es decir, cada solución óptima para el problema minimax, (x^*, y^*), puede expresarse en la forma

$$x^* = \lambda x_1 + (1 - \lambda)x_2,$$
$$y^* = \lambda y_1 + (1 - \lambda)y_2,$$

donde λ es una constante que satisface $0 \le \lambda \le 1$. El valor óptimo de la función objetivo es $c_5/2$.

Ejemplo *10.6*

Considere el ejemplo 10.5 de la Universidad del Lejano Oeste. Como algunos miembros del profesorado tienen alguna discapacidad, el presidente ha decidido ubicar la instalación audiovisual de modo que se minimice la distancia máxima desde la instalación hasta las seis escuelas en el campus. Recuerde que las ubicaciones de las escuelas son

(5, 13),	(8, 18),
(0, 0),	(6, 3),
(14, 20),	(10, 12).

Los valores de las constantes $c_1, \ldots, c_5$ son

$$c_1 = \min_{1 \le i \le n} (a_i + b_i) = \min(18, 26, 0, 9, 34, 22) = 0,$$

$$c_2 = \max_{1 \le i \le n} (a_i + b_i) = \max(18, 26, 0, 9, 34, 22) = 34,$$

$$c_3 = \min_{1 \le i \le n} (-a_i + b_i) = \min(8, 10, 0, -3, 6, 2) = -3,$$

$$c_4 = \max_{1 \le i \le n} (-a_i + b_i) = \max(8, 10, 0, -3, 6, 2) = 10,$$

$$c_5 = \max(c_2 - c_1, c_4 - c_3) = \max(34, 13) = 34.$$

Así, se concluye que

$$x_1 = (c_1 - c_3)/2 = [0 - (-3)]/2 = 1.5,$$

$$y_1 = (c_1 + c_3 + c_5)/2 = (0 - 3 + 34)/2 = 15.5,$$

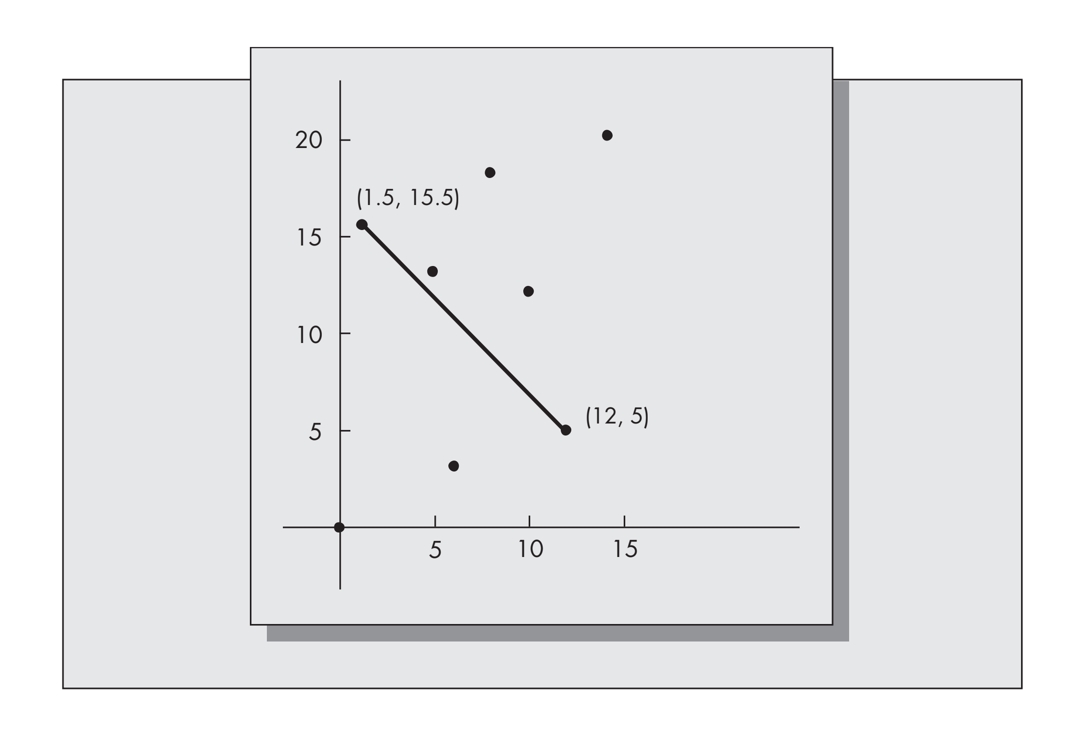
y

$$x_2 = (c_2 - c_4)/2 = (34 - 10)/2 = 12,$$

$$y_2 = (c_2 + c_4 - c_5)/2 = (34 + 10 - 34)/2 = 5.$$

Todos los puntos de la línea que conecta a (x_1, y_1) con (x_2, y_2) son óptimos. El valor de la función objetivo para la(s) solución(es) óptima(s) es 34/2 = 17. Las ubicaciones óptimas del problema minimax se ilustran en la figura 10-30. Recuerde que la solución óptima cuando usamos un objetivo ponderado era (6, 12). Es interesante observar que una solución óptima para este problema, (6, 11), se acerca mucho a esa solución.

FIGURA 10-30
Soluciones óptimas para el objetivo de la ubicación minimax.



Problemas para la sección 10.9

28. Un taller de maquinado tiene cinco máquinas, ubicadas en (3, 3), (3, 7), (8, 4), (12, 3), y (14, 6), respectivamente. Se va a instalar una nueva máquina en el taller con los siguientes números esperados de carga por hora transportada a las máquinas existentes: $\frac{1}{8}$, $\frac{1}{8}$, $\frac{1}{4}$, 1, y $\frac{1}{6}$. El material se transporta a lo largo de corredores paralelos, por lo que resulta apropiada una medida rectilínea de la distancia. Encuentre las coordenadas de la ubicación óptima de la nueva máquina para minimizar la suma ponderada de las distancias rectilíneas desde la nueva máquina hasta las máquinas existentes.
29. Resuelva el problema de ubicación del nuevo centro audiovisual de la Universidad del Lejano Oeste que se describe en el ejemplo 10.5, suponiendo que los pesos son respectivamente 80, 12, 56, 104, 42, y 17 para las escuelas de negocios, educación, ingeniería, humanidades, leyes y ciencias.
30. Armand Bender piensa visitar a seis clientes en Manhattan. Tres están situados en un edificio en la calle 34 y la avenida 7. Los clientes restantes están en la 48 y la 8, la 38 y la 3, y la 42 y la 5. Las calles están separadas 200 pies y las avenidas 400 pies. Él piensa estacionarse una vez y caminar hacia todos los clientes. Suponga que debe regresar a su automóvil después de cada visita para recoger diferentes muestras. ¿En qué ubicación debe estacionarse con objeto de minimizar la distancia total recorrida hacia los clientes? (Sugerencia: al diseñar su cuadrícula, asegúrese de tomar en cuenta el hecho de que las avenidas están separadas una distancia que es el doble de la separación entre calles.)
31. En el problema 30, suponga que el señor Bender puede estacionarse solamente en los cajones ubicados en (1) la calle 40 y la avenida 8, (2) la calle 46 y la avenida 6, y (3) la calle 33 y la avenida 5. ¿Dónde debe planear estacionarse?
32. Un parque industrial consta de 16 edificios. Las corporaciones en el parque están compartiendo el costo de construcción y mantenimiento de un nuevo centro de primeros auxilios. Debido a la distribución del parque, las distancias entre los edificios se aproximan a la medida de la distancia rectilínea. Los pesos para los edificios se determinan con base en la frecuencia de los accidentes. Encuentre la ubicación óptima del centro de primeros auxilios para minimizar la suma ponderada de las distancias rectilíneas hacia los 16 edificios.

Edificio	a_i	b_i	w_i	Edificio	a_i	b_i	w_i
1	0	0	9	9	14	6	11
2	10	3	7	10	19	0	17
3	8	8	4	11	20	4	14
4	12	20	3	12	14	25	6
5	4	9	2	13	3	14	5
6	18	16	12	14	6	6	8
7	4	1	4	15	9	21	15
8	5	3	5	16	10	10	4

33. Dibuje curvas de nivel para el problema de ubicar el taller de máquinas descrito en el problema 28. (Consulte el apéndice 10-B.)
34. Dibuje curvas de nivel para el problema de ubicación descrito en el problema 30. (Consulte el apéndice 10-B.)
35. Dos instalaciones, ubicadas en (0, 0) y (0, 10), tienen pesos respectivos 2 y 1. Dibuje curvas de nivel para este problema. (Consulte el apéndice 10-B.)
36. Resuelva el problema 28 suponiendo un objetivo rectilíneo minimax.
37. Resuelva el problema 30 suponiendo un objetivo rectilíneo minimax.
38. Resuelva el problema 32 suponiendo un objetivo rectilíneo minimax.

10.10 PROBLEMAS DE DISTANCIA EUCLIDIANA

Aunque la medida de distancia rectilínea es apropiada para muchos problemas reales, existen aplicaciones en las cuales la medida apropiada de distancia es la medición en línea recta. Un ejemplo sería ubicar instalaciones de generación de energía con objeto de minimizar la cantidad total de cable eléctrico que debe tenderse para conectar la planta con los clientes. Esta sección considerará el problema euclidiano y una variante que se conoce como el problema de gravedad, que tiene una solución mucho más sencilla.

El problema de gravedad

El problema de gravedad corresponde al caso de un objetivo igual al cuadrado de la distancia euclidiana. Así, el objetivo es encontrar valores de (x, y) para minimizar

$$f(x, y) = \sum_{i=1}^{n} w_i[(x - a_i)^2 + (y - b_i)^2].$$

Este objetivo es apropiado cuando el costo de ubicación de nuevas instalaciones aumenta como una función del cuadrado de la distancia de la nueva instalación hacia las instalaciones existentes. Aunque un objetivo de este tipo no es común, la solución a este problema es directa y frecuentemente se ha usado como una aproximación al problema más común de la distancia en línea recta.

Los valores óptimos de (x, y) se determinan fácilmente por diferenciación. Las derivadas parciales de la función objetivo con respecto a x e y son

$$\frac{\partial f(x, y)}{\partial x} = 2\sum_{i=1}^{n} w_i(x - a_i),$$

$$\frac{\partial f(x, y)}{\partial y} = 2\sum_{i=1}^{n} w_i(y - b_i).$$

Haciendo estas derivadas parciales iguales a cero y despejando los valores de x e y se obtiene la solución óptima

$$x^* = \frac{\sum_{i=1}^{n} w_i a_i}{\sum_{i=1}^{n} w_i},$$

$$y^* = \frac{\sum_{i=1}^{n} w_i b_i}{\sum_{i=1}^{n} w_i}.$$

El término *problema de la gravedad* surge por la siguiente razón. Suponga que se coloca un mapa del área en la cual se va a ubicar la instalación sobre un cartón grueso. Se colocan pesos proporcionales a los números w_i en las ubicaciones de las instalaciones existentes. Entonces la solución de la gravedad es el punto en el mapa respecto al cual todo el conjunto está en equilibrio. (Esta descripción específica es de Keefer, 1934.) Aunque ciertamente podría resolverse el problema de la gravedad de esta manera, es tan fácil encontrar (x^*, y^*) empleando las fórmulas dadas que no hay razón para emplear un modelo físico.

Ejemplo ***10.7*** Encontraremos la solución al problema de ubicar el centro audiovisual de la Universidad del Lejano Oeste suponiendo una medida de ubicación de la distancia euclidiana al cuadrado. Al sustituir los valores de los pesos y de las ubicaciones de los edificios en las fórmulas dadas, obtenemos

$$x^* = 1{,}555/204 = 7.6,$$
$$y^* = 2{,}198/204 = 10.8,$$

que es ligeramente diferente de la solución rectilínea (6, 12).

El problema de la distancia en línea recta

La medición de la distancia en línea recta surge con mucho más frecuencia que la medición de la distancia al cuadrado estudiada en la sección 10.9. El objetivo en este caso es encontrar (x, y) para reducir

$$f(x, y) = \sum_{i=1}^{n} w_i \sqrt{(x - a_i)^2 + (y - b_i)^2}.$$

Desafortunadamente, no es tan fácil encontrar matemáticamente la solución óptima cuando se usa una medida de distancia euclidiana que cuando se usa la distancia euclidiana al cuadrado. Al igual que con el problema de gravedad, también existe un modelo físico que puede construirse para encontrar la solución óptima. En este caso se coloca un mapa del área sobre una mesa. Se perforan orificios en la mesa donde se sitúan instalaciones existentes y se suspenden pesos con hilos a través de los orificios. El tamaño de los pesos debe ser proporcional a los pesos relativos de las ubicaciones. Los hilos se amarran a un anillo. Si no hay fricción, el anillo quedará en reposo en la ubicación de la solución óptima (véase la figura 10-31). Aunque este método puede usarse para encontrar una solución, ciertamente tiene desventajas. En especial, la fricción entre el hilo y la mesa no debe tomarse en cuenta para asegurar que el anillo se detenga en la posición correcta.

La determinación matemática de la solución óptima es más difícil para la distancia euclidiana que para la distancia rectilínea o la distancia euclidiana al cuadrado. No existen soluciones algebraicas simples conocidas; todos los métodos existentes requieren un procedimiento

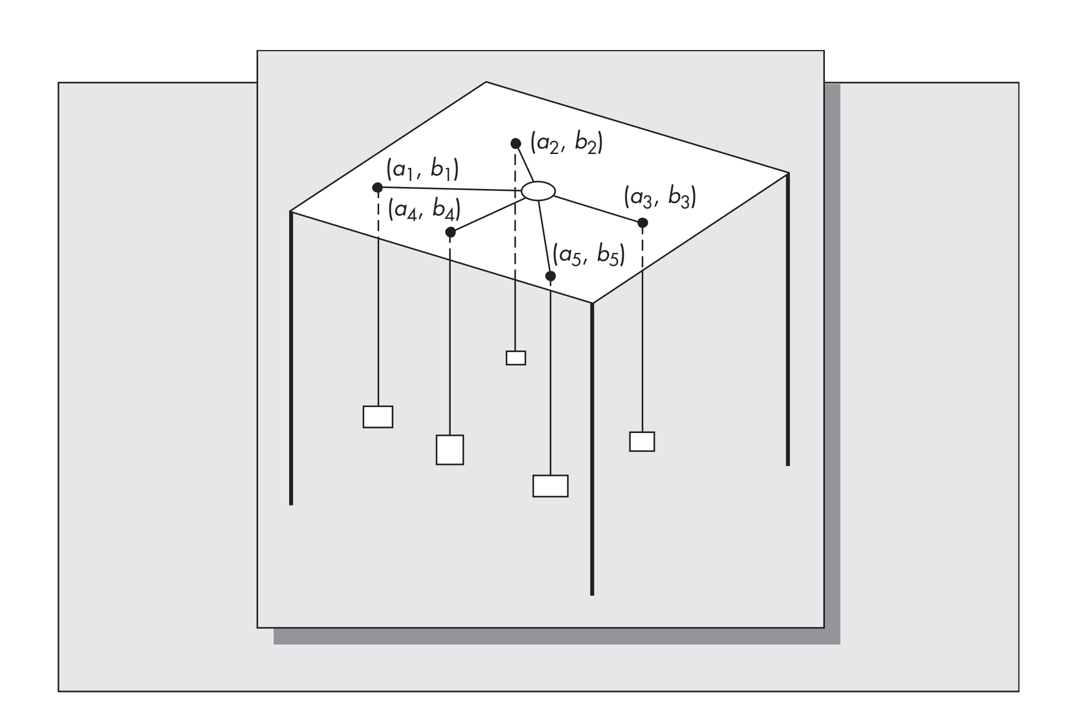


FIGURA 10-31
Solución del problema de la distancia euclidiana mediante el uso de un modelo físico.

iterativo. Describiremos un procedimiento que aparece en Francis y White (1974), el cual arrojará una solución óptima siempre que la ubicación de la nueva instalación no se traslape con la ubicación de una instalación existente.

Defina

$$g_i(x, y) = \frac{w_i}{\sqrt{(x - a_i)^2 + (y - b_i)^2}}.$$

Sea

$$x = \frac{\sum_{i=1}^{n} a_i g_i(x, y)}{\sum_{i=1}^{n} g_i(x, y)},$$

$$y = \frac{\sum_{i=1}^{n} b_i g_i(x, y)}{\sum_{i=1}^{n} g_i(x, y)}.$$

El procedimiento es el siguiente: comience el proceso con una solución inicial (x_0, y_0). Generalmente se recomienda la solución de la gravedad para iniciar el método. Calcule los valores de $g_i(x, y)$ para esta solución y determine nuevos valores de x y y a partir de las fórmulas dadas. Entonces, recalcule $g_i(x, y)$ usando los nuevos valores de x y y, dando lugar a otro par de valores de (x, y). Continúe iterando de esta manera hasta que converjan los valores de las coordenadas. Este procedimiento proporcionará una solución óptima siempre que los valores de (x, y) para cada iteración no correspondan a alguna ubicación existente. El problema es que si sustituimos $x = a_i$ y $y = b_i$, entonces $g_i(x, y)$ no está definido porque el denominador es cero. Es poco probable que ocurra esto, pero si ocurre, se requiere modificar este método (véase Francis y White, 1974).

Ejemplo *10.8*

Resolveremos el problema de ubicación en la universidad suponiendo una medida euclidiana de la distancia. Para iniciar el procedimiento usamos la solución para el problema de la gravedad, la cual es con tres decimales $(x_0, y_0) = (7.622, 10.775)$. La secuencia de valores (x, y) obtenida de la iteración de las ecuaciones dadas para x y y es

$(x_1, y_1) = (7.895, 11.432)$, $(x_5, y_5) = (8.429, 11.880)$,

$(x_2, y_2) = (8.102, 11.715)$, $(x_6, y_6) = (8.481, 11.889)$,

$(x_3, y_3) = (8.251, 11.822)$, $(x_7, y_7) = (8.518, 11.894)$,

$(x_4, y_4) = (8.356, 11.863)$, $(x_8, y_8) = (8.545, 11.899)$.

Es evidente que estos valores comienzan a converger. Al continuar con las iteraciones generalmente se llega a la solución óptima, que es (8.621, 11.907).

Problemas para la sección 10.10

39. Para la situación descrita en el problema 28,

a) Encuentre la solución de gravedad.

b) Use la respuesta obtenida en el inciso *a*) como una solución inicial del problema de la distancia euclidiana. Determine (x_i, y_i) en forma iterativa, para $1 \leq i \leq 5$ si usted está

resolviendo el problema a mano, o para $1 \leq i \leq 20$ si está resolviéndolo con ayuda de una computadora. Con base en sus resultados, estime la solución final.

40. Tres instalaciones existentes se ubican en (0, 0), (5, 5) y (10, 10). Los pesos aplicados a estas instalaciones son 1, 2 y 3, respectivamente. Encuentre la ubicación de una nueva instalación que minimice la distancia euclidiana ponderada hacia las instalaciones existentes.
41. Se instalará un sistema de telecomunicaciones en la ubicación de una compañía. El centro de interruptores que controla el sistema debe ubicarse en uno de los cinco edificios que serán atendidos. La meta es minimizar la cantidad de cableado subterráneo que se requiere. ¿En qué edificio debe ubicarse el centro de interruptores? Suponga una medición en línea recta de la distancia.

Edificio	Ubicación
1	(10, 10)
2	(0, 4)
3	(6, 15)
4	(8, 20)
5	(15, 0)

10.11 OTROS MODELOS DE UBICACIÓN

Los modelos específicos estudiados en las secciones 10.9 y 10.10 han sido denominados como "modelos de ubicación plana" por Francis, McGinnis y White (1983). Como ellos observaron, existen siete supuestos asociados con estos modelos de los problemas de ubicación:

1. Un plano es una aproximación adecuada de una esfera.
2. Cualquier punto del plano es una ubicación válida de una instalación.
3. Las instalaciones pueden idealizarse como puntos.
4. Las distancias entre instalaciones se representan adecuadamente por la distancia planar.
5. Los costos de viaje son proporcionales a la distancia.
6. Ignorando se toman en cuenta los costos fijos.
7. Pueden dejarse de lado los aspectos de distribución.

A menos que las distancias entre las instalaciones sean muy grandes, el supuesto 1) debe ser una aproximación razonable de la realidad. El supuesto 2) puede ser muy restrictivo. En algunas circunstancias, solamente existe un pequeño número de ubicaciones factibles, en cuyo caso el enfoque más sencillo es evaluar el costo de ubicación de la nueva instalación para cada una de estas ubicaciones y seleccionar la que tenga el costo más bajo. Si el número de alternativas factibles es grande, pueden construirse curvas de nivel y considerar solamente aquellas ubicaciones a lo largo de una curva de nivel dada con un aceptable costo bajo.

Según la naturaleza del problema, el supuesto 3) puede ser problemático o no. Por ejemplo, si se usan modelos de ubicación para resolver problemas de distribución en una fábrica, entonces adquiere importancia el tamaño de las instalaciones. Sin embargo, para un problema como la ubicación de almacenes en todo el país, es razonable idealizar las instalaciones como puntos. La hipótesis 4) requiere el uso de una medida específica de distancia para comparar diferentes configuraciones. Sin embargo, una medida de distancia puede no ser suficiente para explicar todas las interacciones de la instalación. Por ejemplo, las distancias rectilíneas se usan comúnmente para problemas en los cuales las instalaciones se ubicarán en las ciudades. Sin embargo, los caminos cerrados, los ríos, o los patrones inusuales de las calles pueden hacer que esta hipótesis sea inexacta.

Los tres supuestos finales se relacionan con aspectos que comúnmente surgen en los problemas de distribución. Los costos de transporte dependen del terreno: por ejemplo, es más caro transportar los bienes por las montañas que por carreteras planas interestatales. Los costos fijos del transporte de bienes pueden ser considerables dependiendo del modo de transporte. Por ejemplo, el costo fijo del transporte de bienes por aire es muy alto.

Existen modelos de ubicación más complejos que enfrentan varias de estas desventajas. Enseguida los revisaremos brevemente.

Localización de instalaciones múltiples

La sección 10.10 trató el problema de ubicar una instalación individual entre n instalaciones existentes. Sin embargo, existen aplicaciones en las cuales la meta es ubicar instalaciones múltiples entre n instalaciones existentes. Por ejemplo, un productor a nivel nacional de productos para el consumidor podría estar considerando dónde ubicar cinco nuevos almacenes regionales.

En algunas circunstancias, los problemas de ubicación de instalaciones múltiples pueden resolverse como una secuencia de problemas de ubicación individual. Es decir, las ubicaciones óptimas de las nuevas instalaciones pueden determinarse una a la vez. Sin embargo, si existe alguna interacción entre las nuevas instalaciones, este enfoque no va a funcionar.

En esta sección se mostrará cómo puede usarse la programación lineal para resolver el problema de la ubicación de instalaciones múltiples mediante la distancia rectilínea. Suponga que las instalaciones existentes tienen ubicaciones en los puntos $(a_1, b_1), \ldots, (a_n, b_n)$ como en la sección 10.10. Asuma que m nuevas instalaciones van a situarse en $(x_1, y_1), \ldots, (x_m, y_m)$. Entonces la función objetivo que debe minimizarse puede escribirse en la forma

$$\text{Minimizar } f_1(\mathbf{x}) + f_2(\mathbf{y}),$$

donde

$$f_1(\mathbf{x}) = \sum_{1\le j<k\le m} v_{jk}|x_j - x_k| + \sum_{j=1}^{m}\sum_{i=1}^{n} w_{ij}|x_j - a_i|$$

y

$$f_2(\mathbf{y}) = \sum_{1\le j<k\le m} v_{jk}|y_j - y_k| + \sum_{j=1}^{m}\sum_{i=1}^{n} w_{ij}|y_j - b_i|.$$

El término v_{jk} mide la interacción de las nuevas instalaciones j y k. Es la presencia de estos términos lo que evita resolver el problema de las instalaciones múltiples como una secuencia de problemas de una instalación individual. Sin embargo, al igual que con el problema de una instalación individual, las ubicaciones óptimas x y y pueden determinarse en forma independiente. Presentamos la formulación de programación lineal para encontrar las coordenadas x óptimas. Las coordenadas óptimas y pueden encontrarse de la misma manera.

La clave de la transformación del problema de encontrar $\mathbf{x}$ para minimizar $f_1(\mathbf{x})$ es eliminar la función del valor absoluto a partir del objetivo, ya que ésta no es estrictamente una función lineal. La manera de hacerlo es un truco normal en la programación lineal. (Se aplicó una técnica similar en la formulación de programación lineal del problema de planeación agregada en la sección 3.5.)

Para cualesquiera constantes a y b escriba $|a - b| = c + d$, pero pida que $cd = 0$ (ya sea c o d o ambos deben ser cero). Si $a > b$, entonces $|a - b| = c$, y si $a < b$, entonces $|a - b| = d$. Así, podemos pensar que c es la parte positiva de $a - b$ y d la parte negativa de $a - b$. Al sustituir $|x_j - x_k| = c_{jk} + d_{jk}$ y $|x_j - a_i| = e_{ij} + f_{ij}$, obtenemos la formulación de programación lineal del problema de la determinación de $\mathbf{x}$:

$$\text{Minimizar } \sum_{1\le j<k\le m} v_{jk}\,(c_{jk} + d_{jk}) + \sum_{j=1}^{m}\sum_{i=1}^{n} w_{ij}\,(e_{ij} + f_{ij})$$

sujeta a

$$x_j - x_k - c_{jk} + d_{jk} = 0, \qquad 1 \le j < k \le n;$$
$$x_j - a_i - e_{ij} + f_{ij} = 0, \qquad 1 \le i \le n, \qquad 1 \le j \le n;$$
$$c_{jk} \ge 0, \quad d_{jk} \ge 0, \qquad 1 \le j < k \le n;$$
$$e_{ij} \ge 0, \quad f_{ij} \ge 0, \qquad 1 \le i \le n, \qquad 1 \le j \le n;$$

x_j sin restricciones en el signo.

No necesitamos incluir explícitamente las restricciones $c_{jk}d_{jk} = 0$ y $e_{ij}f_{ij} = 0$. Puede probarse que para la solución de costo mínimo estas relaciones siempre son válidas (lo cual es ciertamente afortunado ya que éstas no son relaciones lineales). Se requiere una sustitución más antes de resolver el problema. Como los códigos de programación lineal requieren que todas las variables sean no negativas, debemos sustituir $x_j = x_j^+ - x_j^-$, donde $x_j^+ \ge 0$ y $x_j^- \ge 0$. Los códigos comerciales de programación lineal se basan en el método Simplex, el cual es muy eficiente. Pueden resolverse fácilmente en forma realista problemas grandes aun en una computadora personal.

Los problemas de gravedad para instalaciones múltiples requieren la solución de un sistema de ecuaciones lineales, pues también resuelven fácilmente problemas de gravedad que incluyen grandes números de instalaciones. Los problemas euclidianos de instalaciones múltiples se resuelven utilizando una versión multidimensional del método de solución iterativa descrito en la sección 10.10. No se revisarán estos métodos aquí. El lector interesado puede consultar a Francis y White (1974).

Otras extensiones

Instalaciones con áreas positivas

Todos los modelos anteriores suponían que las instalaciones se aproximan mediante puntos en el plano. Si el área de las instalaciones es pequeña comparada con el área cubierta por las ubicaciones disponibles, esta hipótesis es razonable. Sin embargo, al encontrar ubicaciones para máquinas en un taller, las máquinas deben estar lo suficientemente apartadas como para que tengan la capacidad de operar con eficiencia. Tompkins y White (1984) presentan un enfoque que requiere las hipótesis de que las instalaciones tienen forma rectangular y que los pesos tienen una distribución uniforme sobre las áreas. El método se basa en el desarrollo de una analogía entre el problema de la ubicación y el problema de localizar fuerzas en una viga y es similar al procedimiento de construcción de curvas de nivel estudiado en el apéndice 10-B.

Problemas de ubicación-asignación

Frecuentemente la decisión de dónde ubicar las nuevas instalaciones debe acompañarse por la decisión de cuál de las localidades existentes será atendida por cada nueva instalación. Por ejemplo, una compañía puede considerar dónde situar varios almacenes regionales. Además de disponer la ubicación y el número de estos nuevos almacenes, también debe precisarse cuál de los subdistribuidores minoristas será servido por cuáles almacenes.

Los problemas de ubicación-asignación son difíciles de resolver debido al gran número de variables de decisión. Formular la programación matemática del problema, suponiendo que se usa una medida rectangular de distancia, es

$$\text{Minimizar} \sum_{j=1}^{m} \sum_{i=1}^{n} z_{ij} w_{ij} [|x_j - a_i| + |y_j - b_i|] + g(m)$$

sujeta a

$$\sum_{j=1}^{m} z_{ij} = 1 \qquad \text{para } 1 \le i \le n,$$

donde

w_{ij} = Costo por unidad de tiempo por unidad de distancia si la instalación existente i es atendida por la nueva instalación j;

$$z_{ij} = \begin{cases} 1 & \text{si la instalación } i \text{ existente es atendida por la nueva instalación } j, \\ 0 & \text{en otro caso;} \end{cases}$$

m = Número total de nuevas instalaciones, $1 \leq m \leq n$;

(x_j, y_j) = Coordenadas de la instalación existente j, $1 \leq j \leq m$;

(a_i, b_i) = Coordenadas de la instalación existente i, $1 \leq i \leq m$;

$g(m)$ = Costo por unidad de tiempo del suministro de m instalaciones nuevas.

Esta formulación del problema tiene como variables de decisión a m, el número de nuevas instalaciones; z_{ij}, la especificación de cuál de las ubicaciones existentes donde i será servida por la instalación j; y (x_j, y_j), la ubicación de las nuevas instalaciones. La optimización es difícil debido a la presencia de las variables cero-uno z_{ij} e incluir m como una variable de decisión. Comúnmente el problema se resuelve considerando valores sucesivos de $m = 1, 2, \ldots$, y enumerando todas las combinaciones de z_{ij} para cada valor de m. Dados una m y un conjunto de valores z_{ij} fijos, la solución puede obtenerse usando los métodos de ubicación de instalaciones múltiples estudiados anteriormente en esta sección. Sin embargo, el número de valores diferentes de z_{ij} crece rápidamente como una función de m, por lo que de esta manera sólo pueden resolverse problemas de tamaño moderado.

Problemas de ubicación discreta

Los modelos considerados en este capítulo para ubicar nuevas instalaciones suponían que las nuevas instalaciones podían situarse en cualquier parte del plano. Éste no es el caso para la mayoría de las aplicaciones. Las curvas de nivel ayudan a evaluar ubicaciones alternativas, pero no pueden construirse para problemas en los cuales deben localizarse instalaciones múltiples. Un enfoque alternativo es restringir *a priori* las posibles ubicaciones a algún conjunto discreto de posibilidades. Si únicamente hay una sola instalación y el número de ubicaciones posibles es pequeño, el enfoque más sencillo es evaluar el costo de cada ubicación y escoger la más pequeña.

Cuando hay instalaciones múltiples, puede usarse el modelo de asignación estudiado en la sección 10.4 para determinar las ubicaciones óptimas de las nuevas instalaciones. En ciertos tipos de problemas de distribución de almacenes, las nuevas instalaciones pueden ocupar más de un sitio potencial. Por ejemplo, suponga que debemos determinar en qué lugares de un depósito se almacenen k artículos. Suponga que el área de almacenaje apropiada en el depósito se compone de n cuadrados de una cuadrícula y cada artículo almacenado ocupa más de un solo cuadro. Se numeraría cada cuadro y el lugar de almacenaje de un artículo sería especificado por los números de los cuadros de la cuadrícula cubiertos por el artículo. El modelo resultante, estudiado en Francis y White (1974), es una generalización del modelo simple de asignación que se muestra en la sección 10.4. No presentaremos los detalles del modelo aquí.

Modelos de ubicación en redes

Los modelos de ubicación planar suponen que la meta es ubicar una o más instalaciones nuevas con objeto de minimizar alguna función de la distancia que separa a las instalaciones nuevas de las existentes. Generalmente se supone una medida de distancia rectilínea, euclidiana o de otro tipo. Para ciertas aplicaciones, las distancias deben medirse sobre una red existente y no se aproximan con exactitud mediante medidas estándar. El transporte por tierra debe apegarse a las redes de caminos, el transporte por agua debe seguir los carriles de embarque y las rutas marinas, y el transporte por aire se confina a los corredores aéreos previamente determinados. En otras aplicaciones la red puede corresponder a una red de cables de energía o de cables telefónicos. En muchas de estas aplicaciones la nueva instalación o las nuevas instalaciones deben colocarse en, o muy cerca de, una ubicación de la red, y las distancias pueden

medirse solamente en términos de la red. Los modelos de ubicación en red están fuera del alcance de nuestra cobertura. El lector interesado puede consultar los artículos de revisión de Francis, McGinnis y White (1983); y Tansel, Francis y Lowe (1983).

Aspectos internacionales

El problema de ubicar instalaciones es parte del aspecto más amplio de la administración global de las cadenas de abastecimiento. Cada día es más cierto que los negocios están evolucionando hacia corporaciones globales. Éste ya no es el caso únicamente para los gigantes de la industria. Actualmente la globalización juega un papel más importante que nunca antes, y su importancia continuará creciendo. Arntzen y otros (1995) estudian la configuración de la cadena de abastecimiento de la Corporación Digital Equipment. En un caso, ellos muestran cómo se embarcan diferentes partes de computadoras desde Estados Unidos a Europa, de Europa a Brasil y Taiwán, y de Taiwán y China a Europa y de regreso a Estados Unidos. La fabricación puede hacerse en dos o tres países diferentes, y las redes de distribución pueden ser igualmente complejas.

Tradicionalmente los salarios más bajos fueron la principal razón para que las compañías basadas en países desarrollados ubicaran plantas en países menos desarrollados. Ciertamente esto es verdad hoy, General Motors hace gran parte de su ensamblado de automóviles en México. Virtualmente todos los grandes fabricantes de semiconductores tienen instalaciones de fabricación en el extranjero, comúnmente en lugares como Malasia y las Filipinas.

Cohen y Lee (1989) presentan una buena perspectiva de algunos de los aspectos que deben tomarse en cuenta cuando se ubican instalaciones en otros países. Éstos incluyen

1. Los impuestos y los aranceles se basan en los flujos de materiales. Su impacto debe incorporarse a los calendarios de embarque internacional de materiales, el producto intermedio, y el producto terminado.
2. Las tasas de cambio de moneda fluctúan de manera impredecible y afectan los niveles de ganancia en cada país.
3. Las tasas de impuestos corporativos varían considerablemente de un país a otro.
4. El abastecimiento global debe tomar en cuenta los tiempos de demora, los costos, las nuevas tecnologías y la dependencia de países específicos.
5. Las reglas locales de contenidos y las cuotas restringen el flujo de material entre países.
6. Los diseños de producto pueden variar de acuerdo al mercado nacional.
7. Los mecanismos de transferencia de precios deben tomar el lugar del control centralizado.
8. Las diferencias de lenguaje, las normas culturales, la educación y las habilidades deben incorporarse a las decisiones de ubicación.

Sólo recientemente hemos comenzado a entender la complejidad del problema de ubicación de nuevas instalaciones y su efecto sobre las operaciones globales de la cadena de abastecimiento. Los modelos matemáticos bien construidos y bien pensados continuarán ayudándonos a manejar estas redes crecientemente complejas. Sin embargo, muchos de los aspectos abordados en esta sección son difíciles de cuantificar, haciendo así que sea crucial el buen juicio.

Problemas para la sección 10.11

42. Para cada uno de los problemas de ubicación descritos a continuación, comente cuáles de los siete supuestos que se mencionaron en esta sección sean posiblemente violados:

    a) La ubicación de tres máquinas nuevas en un taller de maquinado.
    b) La ubicación de una red internacional de instalaciones de comunicación.
    c) La ubicación de un hospital en un área escasamente poblada.
    d) La ubicación de oficinas de refacciones para apoyar a una organizaron de reparaciones de campo.

43. Considere el problema 28 de este capítulo. Suponga que se instalarán dos nuevas máquinas, A y B, en el taller. La máquina A tiene $\frac{1}{8}$, $\frac{1}{8}$, $\frac{1}{4}$, 1, y $\frac{1}{6}$ como los números esperados de cargas transportadas a las cinco máquinas existentes, respectivamente, y la máquina B tiene $\frac{1}{4}$, $\frac{1}{6}$, 3, $\frac{1}{5}$, y $\frac{1}{2}$ como los números esperados de cargas transportadas, respectivamente. Además, suponga que, en promedio, se transportan dos cargas por hora entre las máquinas nuevas. Suponga una medida de distancia rectilínea.

   a) Formule el problema para determinar las ubicaciones óptimas de las máquinas nuevas como un programa lineal.

   b) Si usted tiene acceso a un código de programación lineal computarizado, resuelva el problema formulado en el inciso *a)*.

44. Considere a la Universidad del Lejano Oeste descrita en el ejemplo 10.5. Suponga que la administración de la universidad ha decidido que se necesitan dos centros audiovisuales. Cada centro tendría instalaciones diferentes. Los números anticipados de los miembros del profesorado que usan cada uno de los centros son

Escuela	Miembros del profesorado que usan el centro A	Miembros del profesorado que usan el centro B
Negocios	13	18
Educación	40	23
Ingeniería	24	17
Humanidades	20	23
Leyes	30	9
Ciencias	16	21

   Además, habrá un total de 16 trabajadores en los centros A y B. Se requerirá que interactúen frecuentemente.

   a) Formule el problema de la determinación de las ubicaciones óptimas de los dos centros audiovisuales como un programa lineal. B. Suponga que siempre se usan distancias rectilíneas.

   b) Si usted tiene acceso a un código computarizado de programación lineal, resuelva el problema formulado en el inciso *a*).

45. Describa los siguientes problemas de ubicación y cómo difieren de aquellos previamente estudiados en este capítulo:

   a) Problemas de ubicación-asignación

   b) Problemas discretos de ubicación

   c) Problemas de ubicación en redes

10.12 NOTAS HISTÓRICAS

Los problemas estudiados en este capítulo tienen una larga historia. La determinación de distribuciones adecuadas para las instalaciones de producción es un problema que se remonta al inicio de la Revolución Industrial, aunque es evidente que el desarrollo de técnicas analíticas para encontrar las distribuciones es reciente. Parece haberse publicado poco en relación con los métodos analíticos de distribución antes de 1950. Apple (1977) enlista varios libros que tratan el problema del diseño de plantas, publicados en los inicios de la década de los años cincuenta. Las técnicas computarizadas de distribución estudiadas en este capítulo se desarrollaron en las décadas de los sesenta y setenta. CRAFT, uno de los primeros métodos computarizados y que aún es muy popular en la actualidad, es de Buffa, Armour y Vollmann (1964). ALDEP es de Seehof y Evans (1967), CORELAP de Lee y Moore (1967), COFAD de Tompkins y Reed

(1976), y PLANET de Deisenroth y Apple (1972). Tanto ALDEP como CRAFT están disponibles en la Corporación IBM como parte de su biblioteca SHARE.

Algunos de los problemas de ubicación estudiados en este capítulo se remontan a cientos de años. El problema de encontrar la ubicación de una instalación nueva individual para minimizar la suma de las distancias euclidianas a las instalaciones existentes se ha denominado como el problema de Steiner-Weber o el problema general de Fermat. Francis y White (1974) afirman que el problema con exactamente tres instalaciones fue propuesto por Fermat y fue resuelto por el matemático Torricelli antes de 1640. El trabajo sobre el problema de la distancia rectilínea es relativamente reciente y fue instigado por un artículo de Hakimi (1964). Actualmente continúa la investigación sobre el descubrimiento de técnicas eficientes de solución para ubicar instalaciones múltiples usando diferentes medidas de distancia.

10.13 Resumen

Este capítulo trató dos problemas importantes en la logística: la distribución más eficiente de las instalaciones y la mejor ubicación de las instalaciones nuevas en relación con las ya existentes. En cierto sentido, el problema de la distribución es un tipo especial de problema de ubicación, porque la meta es encontrar la mejor ubicación de las instalaciones dentro de un límite específico.

Los métodos analíticos para la distribución estudiados en este capítulo suponen que el objetivo es minimizar alguna función de la distancia que separa a las instalaciones. Este punto de vista es probablemente el más apropiado para los problemas de diseño de planta y el menos conveniente para otros problemas en los cuales los factores cualitativos juegan un papel mayor. Dos diagramas son importantes para el análisis de las distribuciones: el *diagrama de relación de actividades* (abreviado como diagrama de relación) y el *diagrama de-hacia*. Con objeto de construir una gráfica de relación, a cada par de instalaciones se le da un código de letras *A* (absolutamente necesario), *E* (especialmente importante), *I* (importante), *O* (importancia ordinaria), *U* (sin importancia) o *X* (indeseable), que representan la conveniencia de ubicar instalaciones cercanas entre sí. Un diagrama de-hacia puede especificar la distancia entre pares de instalaciones, los números de viajes de manejo de material por unidad de tiempo entre dos instalaciones, o el costo de los viajes de manejo de material entre dos instalaciones. Tanto los diagramas de relación como los diagramas de-hacia son útiles para evaluar la calidad de una distribución.

En la sección 10.3 se analizan los tipos de distribuciones. Los dos tipos más comunes son las *distribuciones por producto* y las *distribuciones por proceso*. La distribución por producto generalmente es una línea de transferencia fija arreglada según la secuencia de los pasos de manufactura que se requieren. Una distribución por proceso agrupa a las máquinas con funciones similares. El ruteo de las partes varía de un producto a otro. La distribución por producto es apropiada para un ambiente de taller de tareas de bajo volumen. *Las distribuciones por posición fija* se usan para productos que son demasiado grandes para moverlos. Recientemente ha habido un interés considerable en las distribuciones basadas en la *tecnología de grupos*. Las partes se agrupan en familias, y se desarrollan células de máquinas consistentes con este agrupamiento. Las distribuciones de tecnología de grupos son apropiadas para las fábricas automatizadas.

El *modelo de asignación* puede usarse para solucionar problemas relativamente sencillos de distribución y de ubicación. Con objeto de usar el algoritmo de asignación, suponemos que por cada colocación de una máquina (digamos) en una ubicación, podemos evaluar el costo de esa asignación. Esto supone que no hay interacción entre las máquinas. Si se presenta una interacción, existe una formulación de asignación cuadrática, pero los modelos de *asignación cuadrática* son mucho más difíciles de resolver que los modelos de asignación simple.

Estudiamos cinco técnicas *computarizadas de distribución*: CRAFT, COFAD, ALDEP, PLANET y CORELAP. Tanto CRAFT como COFAD son rutinas de mejora. Esto significa que ambas requieren que el usuario especifique una distribución inicial. El programa prosigue para considerar el intercambio de los pares adyacentes de las instalaciones con objeto de lograr una mejora. Por otro lado, ALDEP, PLANET y CORELAP son rutinas de construcción que cons-

truyen la distribución desde cero. Como frecuentemente las rutinas de construcción conducen a departamentos con formas complejas, generalmente se prefieren las rutinas de mejora. También observamos que existe un alud de nuevos productos de software actualmente disponibles para la computadora personal y para las estaciones de trabajo UNIX. Muchos de estos productos se basan en programas de dibujo con librerías grandes de iconos gráficos. Estos productos usan métodos conceptualmente similares a los programas antes mencionados, pero son mucho más amigables con el usuario.

El capítulo incluyó un análisis de los *sistemas flexibles de manufactura* (FMS). Un FMS es una colección de máquinas enlazadas por un sistema automatizado de manejo de material que generalmente está controlado por una computadora central. Los FMS se usan principalmente en las industrias donde se trabajan metales y son una selección apropiada cuando se requieren un volumen de medio a grande y una variedad moderada de tipos de parte. La desventaja de estos sistemas es el costo, que puede ser tan alto como $10 millones o más. Las células flexibles de manufactura son una alternativa de costo más bajo a escala reducida.

La segunda parte del capítulo se ocupó de los métodos de ubicación de las nuevas instalaciones. Los *modelos de ubicación* son apropiados cuando se ubican una o más instalaciones nuevas dentro de un área específica que ya contiene un número finito de instalaciones existentes. El objetivo es ubicar nuevas instalaciones para minimizar alguna función de la distancia que separa a las instalaciones nuevas de las existentes. Se consideraron tres medidas de distancia: rectilínea, euclidiana y euclidiana cuadrada. Las dos primeras son las más comunes para describir problemas reales y dependen de si hay movimiento de acuerdo con un patrón reticular de las calles (rectilíneo) o se mide mediante distancias en línea recta (euclidiana).

La solución óptima al problema de la distancia rectilínea ponderada es ubicar las coordenadas (x, y) en la *mediana* de las coordenadas existentes. Si se usa una medición de distancia euclidiana al cuadrado, la ubicación óptima de la nueva instalación está en el centro de gravedad de las coordenadas existentes. No se conoce ninguna solución algebraica simple para el problema de la distancia euclidiana, pero existen técnicas de solución iterativa. El capítulo también incluye una breve discusión sobre varios problemas de ubicación más complejos, incluyendo la ubicación de instalaciones múltiples, la ubicación de instalaciones que tengan áreas diferentes de ceros, problemas de ubicación-asignación, problemas de ubicación discreta y modelos de ubicación en redes. Finalmente, el capítulo concluyó con un análisis en torno a temas de interés al ubicar nuevas instalaciones en otros países. A medida que aumente la globalización de las redes en cadena de manufactura y de abastecimiento, estos aspectos jugarán un papel cada vez más importante en la planeación de las ubicaciones.

Problemas adicionales sobre distribución y ubicación

46. Una compañía de bienes raíces desea abrir cuatro oficinas nuevas en el área de Boston. Se dispone de seis sitios potenciales. Con base en el número de empleados en cada oficina y en la ubicación de las propiedades que cada empleado debe administrar, la compañía estimó el tiempo total de viaje en horas por día para cada oficina y para cada ubicación. Encuentre la asignación óptima de las oficinas a los sitios para minimizar el tiempo de viaje del empleado.

		Oficinas			
		A	**B**	**C**	**D**
Sitios	**1**	10	3	3	8
	2	13	5	2	6
	3	12	9	9	4
	4	14	2	7	7
	5	17	7	4	3
	6	12	8	5	5

47. Una gran cadena de supermercados en el sureste requiere cinco almacenes adicionales en el área de Atlanta. Se han identificado cinco sitios para estos almacenes. Los costos anuales del transporte (en $000) para cada almacén se dan en la siguiente tabla. Encuentre la asignación de los almacenes a los sitios para minimizar los costos anuales totales del transporte.

		Almacenes				
		A	**B**	**C**	**D**	**E**
	1	41	47	38	46	50
	2	39	37	42	36	45
Sitios	**3**	43	46	45	42	46
	4	51	54	47	58	56
	5	44	40	42	41	45

48. Un taller de maquinado que se localiza en los suburbios de Los Ángeles acepta órdenes a la medida de un número de compañías de alta tecnología en el sureste de California. El taller de maquinado consiste de cuatro departamentos: A (tornos), B (taladros), C (esmeriladoras) y D (lijadoras). El diagrama de-hacia que muestra las distancias en pies entre los centros de los departamentos se muestra a continuación.

		HACIA			
		A	**B**	**C**	**D**
	A		45	63	32
D	**B**	29		27	46
E	**C**	63	75		68
	D	40	30	68	

El taller ha aceptado órdenes para generar cuatro productos: P1, P2, P3 y P4. El ruteo para manufacturar estos productos y las tasas semanales de producción son

Producto	Ruteo	Producción semanal
P1	A–B–C–D	200
P2	A–C–D	600
P3	B–D	400
P4	B–C–D	500

Suponga que los productos se producen en lotes de tamaño 25.

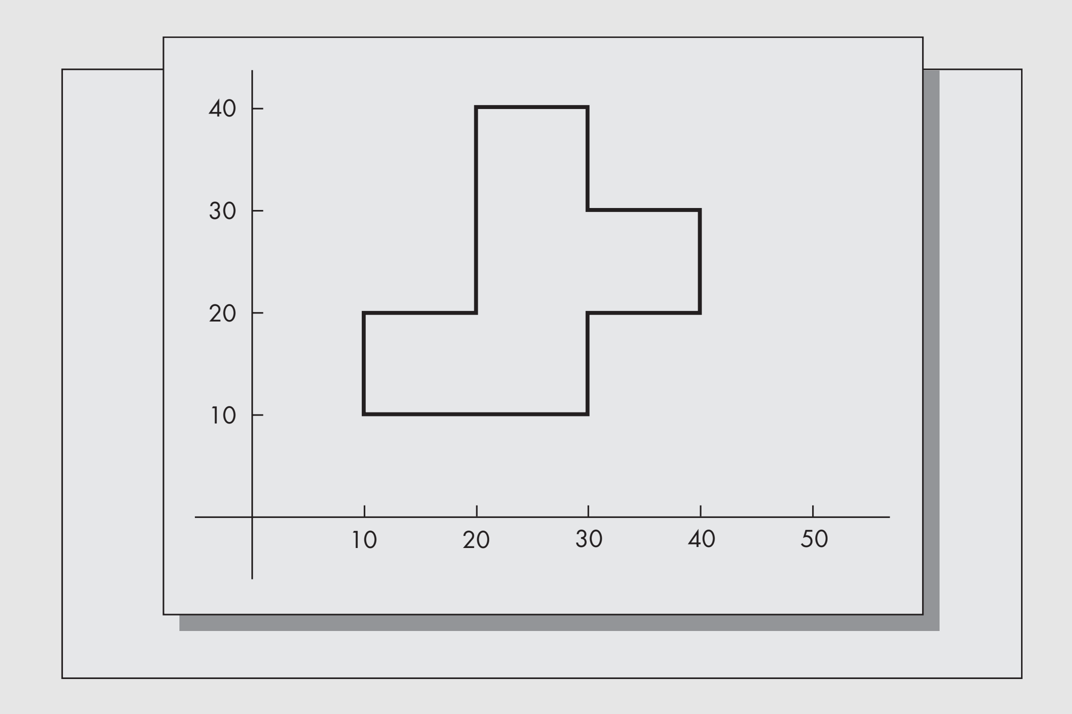
a) Convierta esta información a un diagrama de-hacia que dé los números de viajes de manejo de material por semana entre departamentos.

b) Si se estima que el costo de transporte de un lote por pie es de $1.50, convierta el diagrama de-hacia que encontró en el inciso *a*) a uno que dé el costo de manejo de material por semana entre departamentos.

c) Desarrolle un diagrama de relación de actividades para estos cuatro departamentos con base en los resultados del inciso *b*). Suponga que A se asigna al costo más alto y *O* al más bajo, con las jerarquizaciones *E* e *I* asignadas a los costos que se sitúan entre los extremos.

d) Suponga que el taller de maquinado se ubica en un edificio que mide 60 pies por 80 pies. Además, suponga que los departamentos tienen forma rectangular con las siguientes dimensiones:

FIGURA 10-32
Forma de la instalación (para el problema 49).

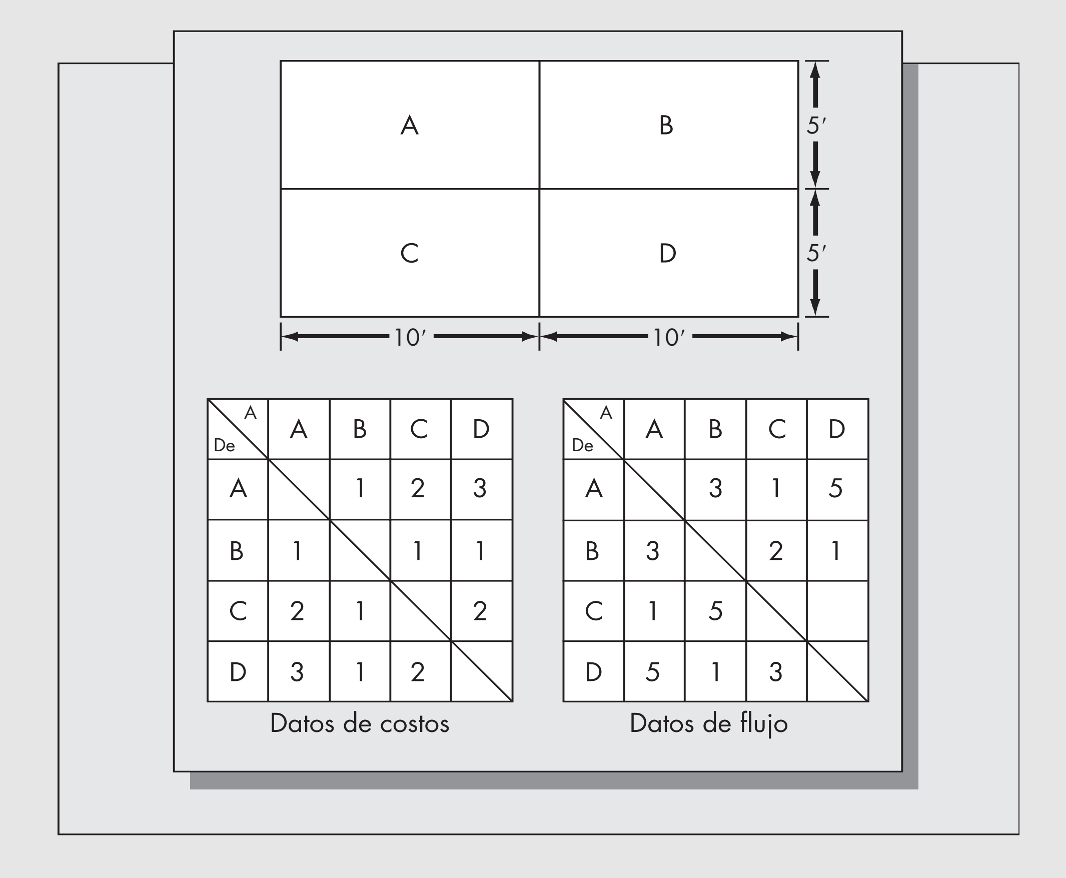


Departamento	Dimensiones
A	20 × 30
B	40 × 20
C	45 × 55
D	37 × 25

Dibuje una distribución consistente con la gráfica de relación que obtuvo en el inciso *c*).

49. Una instalación tiene la forma mostrada en la figura 10-32. Usando los métodos descritos en el apéndice 10-A, encuentre dónde se ubica el centroide de esta instalación.
50. En la figura 10-33 se muestra una instalación inicial para cuatro departamentos y un diagrama de datos de-hacia. Utilizando la técnica de CRAFT de intercambio por pares, encuentre la distribución recomendada para minimizar los costos totales de manejo de materiales.
51. En la figura 10-34 se muestra una instalación inicial para cinco departamentos y un diagrama de datos de-hacia. Suponiendo que los departamentos A y D están en ubicaciones fijas y no pueden moverse, encuentre la distribución recomendada por CRAFT para los departamentos B, C y E. Suponga que el objetivo es minimizar la distancia total recorrida.
52. Considere la gráfica de relación para el restaurante de comida rápida Meat Me, dada en la figura 10-2. Suponga que las áreas requeridas para cada departamento son

Departamento	Área requerida (pies cuadrados)
Cocinado de hamburguesas	200
Cocinado de papas fritas	200
Empaque y almacenaje	100
Despachadores de refrescos	100
Servidores de mostrador	400
Servidor para automovilistas	100



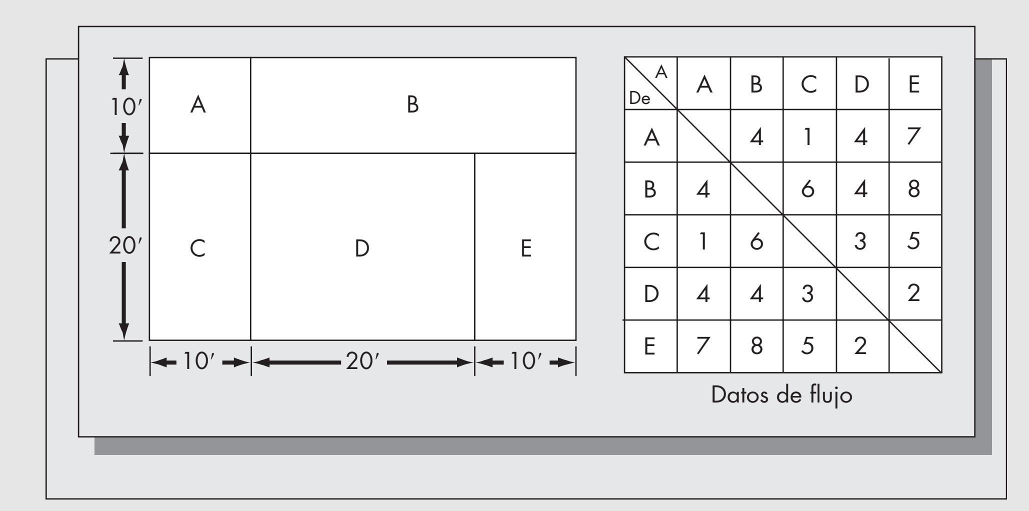
De \ A	A	B	C	D
A		1	2	3
B	1		1	1
C	2	1		2
D	3	1	2	

Datos de costos

De \ A	A	B	C	D
A		3	1	5
B	3		2	1
C	1	5		
D	5	1	3	

Datos de flujo

FIGURA 10-33
Distribución y diagramas de-hacia (para el problema 50).



De \ A	A	B	C	D	E
A		4	1	4	7
B	4		6	4	8
C	1	6		3	5
D	4	4	3		2
E	7	8	5	2	

Datos de flujo

FIGURA 10-34
Distribución y diagrama de-hacia (para el problema 51).

Suponga un ancho de barrido de 2 cuadros y dimensiones de la instalación de 5 por 9 cuadros, donde cada cuadro tiene 5 pies por lado. Como resultado, por ejemplo, el departamento de cocinado de hamburguesas requiere 8 cuadros. Use el enfoque ALDEP para desarrollar una distribución para estructurar el restaurante. Comente qué tan prácticos son sus resultados.

53. Frank Green, un técnico independiente de reparación de TV, está considerando comprar una casa en Ames, Iowa, que usará como base de operaciones para su negocio de reparaciones. Las fuentes de negocios primarias de Frank son 10 cuentas industriales ubicadas en el área de Ames. Él ha superpuesto una cuadrícula sobre un mapa de la ciudad y ha

determinado las siguientes ubicaciones para estos clientes, así como el número esperado de llamadas que recibe al mes:

Cliente	Ubicación en la cuadrícula	Llamadas esperadas por mes
1	(5, 8)	2
2	(10, 3)	1
3	(14, 14)	1
4	(2, 2)	3
5	(1, 17)	1
6	(18, 25)	$\frac{1}{2}$
7	(14, 3)	$\frac{1}{4}$
8	(25, 4)	4
9	(35, 1)	3
10	(16, 21)	$\frac{1}{6}$

Encuentre la ubicación óptima de su casa, suponiendo

a) Una medida de distancia rectilínea ponderada.

b) Una medida de distancia euclidiana cuadrada.

c) La meta es minimizar la distancia rectilínea máxima hasta cualquier cliente.

54. Una compañía de electrónica ubicada cerca de Phoenix, Arizona, está considerando dónde ubicar un nuevo conmutador telefónico que enlace a cinco edificios. Los edificios se ubican en (0, 0), (2, 6), (10, 2), (3, 9) y (0, 4). El objetivo es ubicar el conmutador para minimizar el cableado requerido para esos cinco edificios.

a) Determine la solución de gravedad.

b) Determine la ubicación óptima suponiendo una medición de distancia en línea recta. (Si usted está resolviendo este problema a mano, itere las ecuaciones apropiadas cuando menos cinco veces y estime la solución óptima.)

55. Considere tres ubicaciones en (0, 0), (0, 6) y (3, 3) con pesos iguales. Usando los métodos descritos en el apéndice 10-B, encuentre las curvas de nivel para el problema de ubicación rectilínea.

56. Una compañía está considerando dónde ubicar su cafetería para dar servicio a seis edificios. Las ubicaciones de los edificios y la fracción de los empleados de la compañía que trabajan en estas ubicaciones son

Edificio	a_i	b_i	Fracción de la fuerza de trabajo
A	2	6	$\frac{1}{12}$
B	1	0	$\frac{1}{12}$
C	3	3	$\frac{1}{6}$
D	5	9	$\frac{1}{4}$
E	4	2	$\frac{1}{4}$
F	10	7	$\frac{1}{6}$

a) Encuentre la ubicación óptima de la cafetería para minimizar la distancia rectilínea ponderada hacia todos los edificios.

b) Encuentre la ubicación óptima de la cafetería para minimizar la distancia rectilínea máxima hacia todos los edificios.

c) Encuentre la solución de gravedad.

d) Suponga que la cafetería debe ubicarse en uno de los edificios. ¿En cuál edificio deberá ubicarse si la meta es reducir la distancia rectilínea ponderada?

e) Resuelva el inciso *d* suponiendo una distancia euclidiana ponderada.

Problemas en hoja de cálculo para el capítulo 10



57. Diseñe una hoja de cálculo para calcular la distancia rectilínea total desde un conjunto de hasta 10 ubicaciones existentes hasta cualquier otra ubicación. Suponga que las ubicaciones existentes se colocan en las columnas A y B y la nueva ubicación en las celdas D1 y E1. Inicie la columna A con el valor de la celda D1 y la columna B con el valor de la celda E1, de modo que la distancia total se calcule correctamente cuando haya menos de 10 ubicaciones.

a) Suponga que las instalaciones existentes se ubican en (0, 0), (5, 15), (110, 120), (35, 25), (80, 10), (75, 20), (8, 38), (50, 65), (22, 95) y (44, 70), y la nueva instalación debe ubicarse en (50, 50). Determine la distancia rectilínea total de la nueva instalación hacia las instalaciones existentes.

b) Suponga que la nueva instalación del inciso *a*) puede ubicarse solamente en $x = 0, 5, 10, \ldots, 100$ y $y = 0, 10, 20, \ldots, 100$. Variando sistemáticamente las coordenadas x y y, encuentre la ubicación óptima de la nueva instalación.



58. Resuelva el problema 57, suponiendo una medición de distancia euclidiana.



59. Resuelva el problema 32 usando una hoja de cálculo electrónica. Para hacerlo, introduzca los números de los edificios en la columna A, las coordenadas x en la columna B, y los pesos asociados en la columna C. Organice las columnas A, B y C en orden ascendente usando la columna B como la clave primaria para la ordenación. Ahora acumule los pesos (columna C) en la columna D usando la función de suma. Divida el peso total acumulado entre 2 e identifique visualmente el valor óptimo de la coordenada x. Será donde el peso acumulado sobrepase por primera vez a la mitad del peso total acumulado. Repita el proceso para las coordenadas y.



60. Diseñe una hoja de cálculo electrónica para calcular la solución óptima al problema de la gravedad. Permita hasta 20 ubicaciones. Sea la columna A el número de ubicación, la columna B las coordenadas x $(a_1, \ldots, a_n)$ de las ubicaciones existentes, y la columna C las coordenadas y $(b_1, \ldots, b_n)$ de las ubicaciones existentes. Almacene la solución óptima en la celda D1. Encuentre la solución de gravedad para el problema 32 usando su hoja de cálculo.



61. Amplíe los resultados del problema 60 para encontrar la ubicación óptima de una nueva instalación entre un conjunto de instalaciones existentes suponiendo una medida de distancia euclidiana en línea recta. Sea la columna E que corresponde a $g_i(x, y)$ donde los valores iniciales de (x, y) aparecen en F1 y F2. En las ubicaciones G1 y G2, almacene

$$x = \frac{\sum_{i=1}^{n} a_i g_i(x, y)}{\sum_{i=1}^{n} g_i(x, y)},$$

$$y = \frac{\sum_{i=1}^{n} b_i g_i(x, y)}{\sum_{i=1}^{n} g_i(x, y)}.$$

Empiece con la solución de gravedad en las celdas F1 y F2. Después del cálculo, remplace los valores en las celdas F1 y F2 con los valores en las celdas G1 y G2. Continúe de esta manera hasta que la solución converja. Usando su hoja de cálculo,

a) Resuelva el problema 32.
b) Resuelva el problema 41.
c) Resuelva el problema 53 suponiendo una medida de distancia euclidiana.
d) Resuelva el problema 56 suponiendo una medida de distancia euclidiana.

Apéndice 10-A

Determinación de centroides

El centroide de cualquier objeto es otro término para las coordenadas físicas del centro de gravedad. Para una placa de densidad uniforme, sería el punto donde la placa estaría en un equilibrio exacto. Sea R cualquier región en el plano. El centroide de R se define mediante dos puntos $\bar{x}$, $\bar{y}$. Para encontrar estos dos puntos, primero debemos obtener los momentos de R, M_x, y M_y, que están dados por la fórmulas

$$M_x = \int_R\int x\,dx\,dy,$$
$$M_y = \int_R\int y\,dx\,dy.$$

Sea $A(R)$ el área de R. Entonces el centroide de R está dado por

$$\bar{x} = \frac{M_x}{A(R)}, \qquad \bar{y} = \frac{M_y}{A(R)}.$$

Ahora obtenemos expresiones explícitas para los momentos donde R es una suma finita de rectángulos. Suponga que R es un rectángulo simple como se ilustra en la figura 10-35. Entonces

$$M_x = \int_{y_1}^{y_2} dy \int_{x_1}^{x_2} x\,dx = \int_{y_1}^{y_2} dy \frac{x^2}{2}\bigg|_{x_1}^{x_2}$$
$$= \int_{y_1}^{y_2} \frac{dy\left(x_2^2 - x_1^2\right)}{2} = \frac{x_2^2 - x_1^2}{2}(y_2 - y_1),$$
$$M_y = \int_{x_1}^{x_2} dx \int_{y_1}^{y_2} y\,dy = \int_{x_1}^{x_2} dx \frac{y^2}{2}\bigg|_{y_1}^{y_2}$$
$$= \int_{x_1}^{x_2} \frac{dx\left(y_2^2 - y_1^2\right)}{2} = \frac{y_2^2 - y_1^2}{2}(x_2 - x_1).$$

Observe que como el área del rectángulo es $(x_2 - x_1)(y_2 - y_1)$, y $x_2^2 - x_1^2 = (x_2 - x_1)(x_1 + x_2)$ (y en forma similar para $y_1^2 - y_2^2$), obtenemos

$$\bar{x} = \frac{(x_1 + x_2)}{2}, \qquad \bar{y} = \frac{(y_1 + y_2)}{2}.$$

Las fórmulas para los momentos de un rectángulo pueden usarse para encontrar el centroide si R también consiste de un grupo de rectángulos. Suponga que R puede subdividirse en k

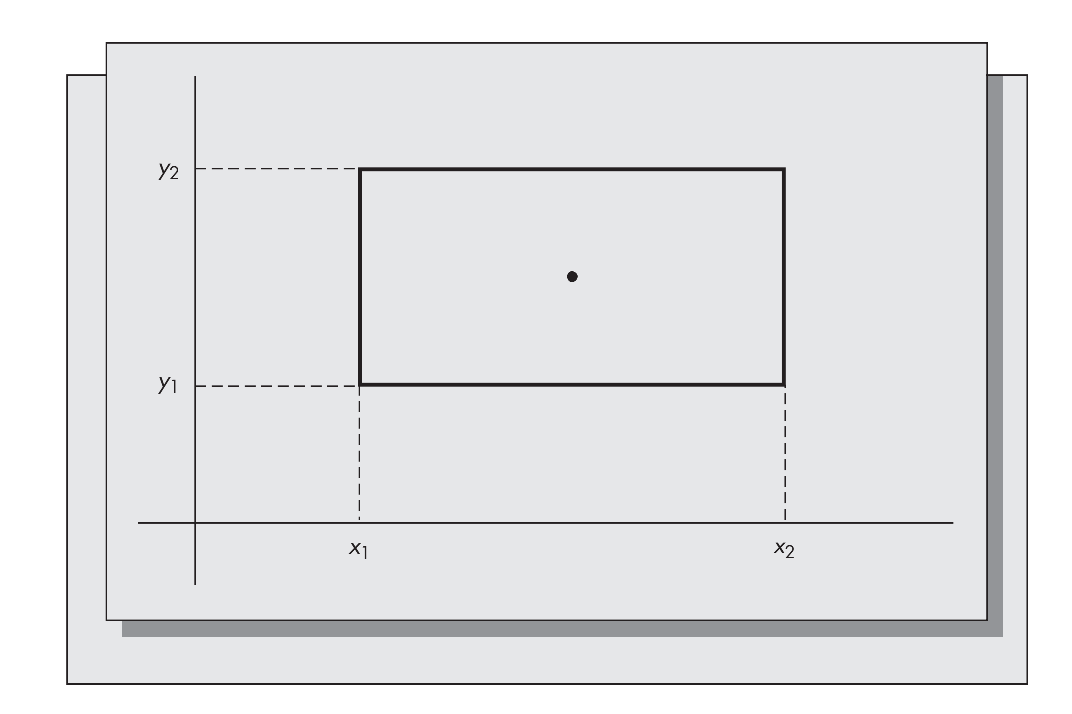


FIGURA 10-35
El centroide de un rectángulo.

rectángulos etiquetados $R_1, R_2, \ldots, R_k$ con los respectivos límites definidos por $[(x_{1i}, x_{2i}), (y_{1i}, y_{2i})]$ para $1 \leq i \leq k$. Como

$$M_x = \int_R \int x\, dx\, dy = \sum_{i=1}^{k} \int_{R_i} \int x\, dx\, dy,$$

se concluye que

$$M_x = \sum_{i=1}^{k} \frac{{x_{2i}}^2 - {x_{1i}}^2}{2}(y_{2i} - y_{1i}).$$

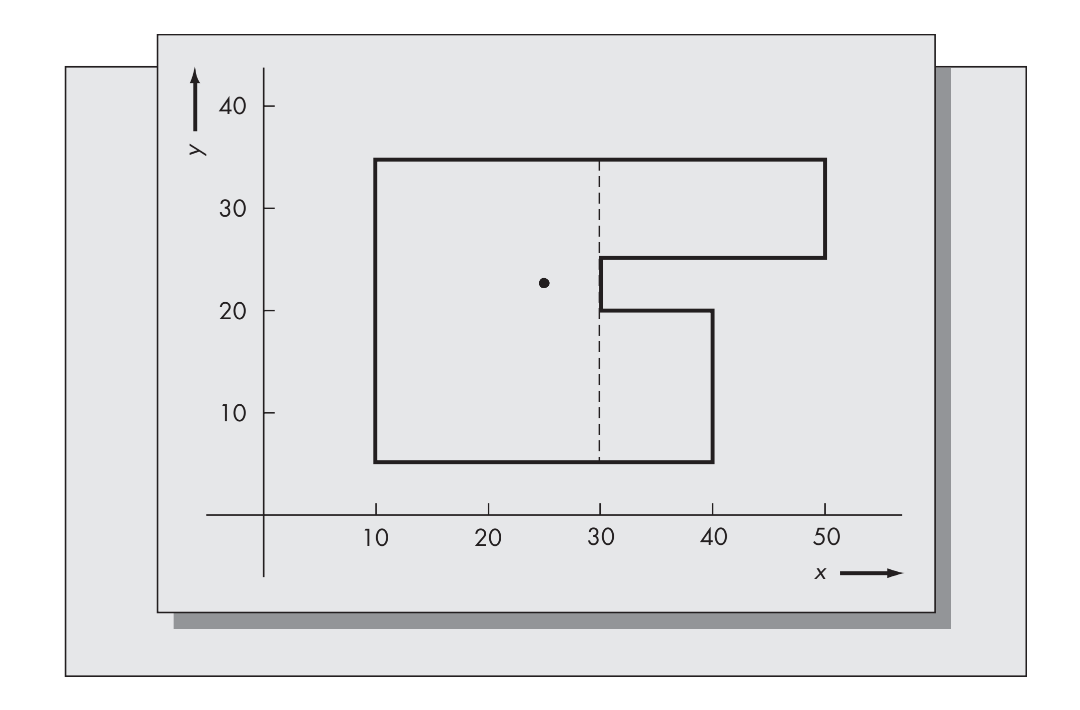
Similarmente,

$$M_y = \sum_{i=1}^{k} \frac{{y_{2i}}^2 - {y_{1i}}^2}{2}(x_{2i} - x_{1i}).$$

Ejemplo *10A.1* Considere la región R ilustrada en la figura 10-36. Determinaremos el centroide de la región usando las fórmulas dadas. La región puede descomponerse en tres rectángulos de varias maneras diferentes. Para la que se ilustra en la figura 10-36, tenemos

$$\begin{aligned} x_{11} &= 10, & x_{21} &= 30, \\ y_{11} &= 5, & y_{21} &= 35, \\ x_{12} &= 30, & x_{22} &= 40, \\ y_{12} &= 5, & y_{22} &= 20, \\ x_{13} &= 30, & x_{23} &= 50, \\ y_{13} &= 25, & y_{23} &= 35. \end{aligned}$$

FIGURA 10-36
Centroide de una figura compuesta por rectángulos.



Sustituyendo en las fórmulas dadas, obtenemos

$$M_x = \frac{30^2 - 10^2}{2}(35 - 5) + \frac{40^2 - 30^2}{2}(20 - 5) + \frac{50^2 - 30^2}{2}(35 - 25)$$

$$= (400)(30) + (350)(15) + (800)(10) = 25\,250,$$

$$M_y = \frac{35^2 - 5^2}{2}(30 - 10) + \frac{20^2 - 5^2}{2}(40 - 30) + \frac{35^2 - 25^2}{2}(50 - 30)$$

$$= (600)(20) + (187.5)(10) + (300)(20) = 19\,875.$$

El área total de R es

$$A(R) = (20)(30) + (10)(15) + (20)(10) = 950.$$

Se concluye que el centroide es

$$\bar{x} = \frac{25\,250}{950} = 26.579,$$

$$\bar{y} = \frac{19\,875}{950} = 20.921.$$

El centroide está marcado por un punto en la figura 10-36.

Apéndice 10-B

Cálculo de curvas de nivel

Este apéndice esboza el procedimiento para calcular curvas de nivel, o líneas de isocosto, tales como las que se ilustran en las figuras 10-28 y 10-29. La justificación teórica para este procedimiento aparece en Francis y White (1974).

1. Grafique los puntos (a_1, b_1), (a_2, b_2), . . . , (a_n, b_n) en papel gráfico. Dibuje una línea horizontal (paralela al eje x) y una línea vertical (paralela al eje y) a través de cada punto.

2. Numere las líneas horizontal y vertical en secuencia de izquierda a derecha y de arriba a abajo. (Si ninguno de los puntos originales es colineal, habrá exactamente n líneas horizontales y n verticales.)

3. Sea C_i la suma de los pesos asociados con los puntos a lo largo de la línea vertical j, y D_i la suma de los pesos asociados con los puntos a lo largo de la línea horizontal i.

4. Calcule los siguientes números:

$$M_0 = -\sum_{i=1}^{n} w_i, \qquad N_0 = M_0 = -\sum_{i=1}^{n} w_i,$$

$$M_1 = M_0 + 2C_1, \qquad N_1 = N_0 + 2D_1,$$

$$M_2 = M_1 + 2C_2, \qquad N_2 = N_1 + 2D_2,$$

etcétera.

(Los valores finales de M_i y N_j serán $+\sum_{i=1}^{n} w_i$.)

5. Defina la región (i, j) como la región limitada por las líneas verticales i-ésima e $(i + 1)$-ésima y las líneas horizontales j-ésima y $(j + 1)$-ésima. Las regiones a la izquierda de la primera línea vertical se etiquetan como $(0, j)$, y aquellas debajo de la primera línea horizontal se etiquetan como $(i, 0)$. La pendiente de cualquier curva de nivel que pasa por la región (i, j) está dada por

$$S_{i,j} = -M_i/N_j.$$

Una vez que se determinan las pendientes, se construye una curva de nivel iniciando en cualquier punto y moviéndose hacia cada región para el ángulo determinado por la pendiente calculada en el paso 5. A continuación presentamos un ejemplo simple para ilustrar el método.

Ejemplo *10B.1* Suponga $(a_1, b_1) = (1, 1)$, $(a_2, b_2) = (5, 3)$, y $(a_3, b_3) = (3, 8)$, y $w_1 = 2$, $w_2 = 3$, y $w_3 = 6$. El primer paso es graficar estos puntos en una cuadrícula como lo hemos hecho en la figura 10-37. El trazo de líneas verticales y horizontales a través de cada uno de los tres puntos resulta en tres líneas verticales etiquetadas

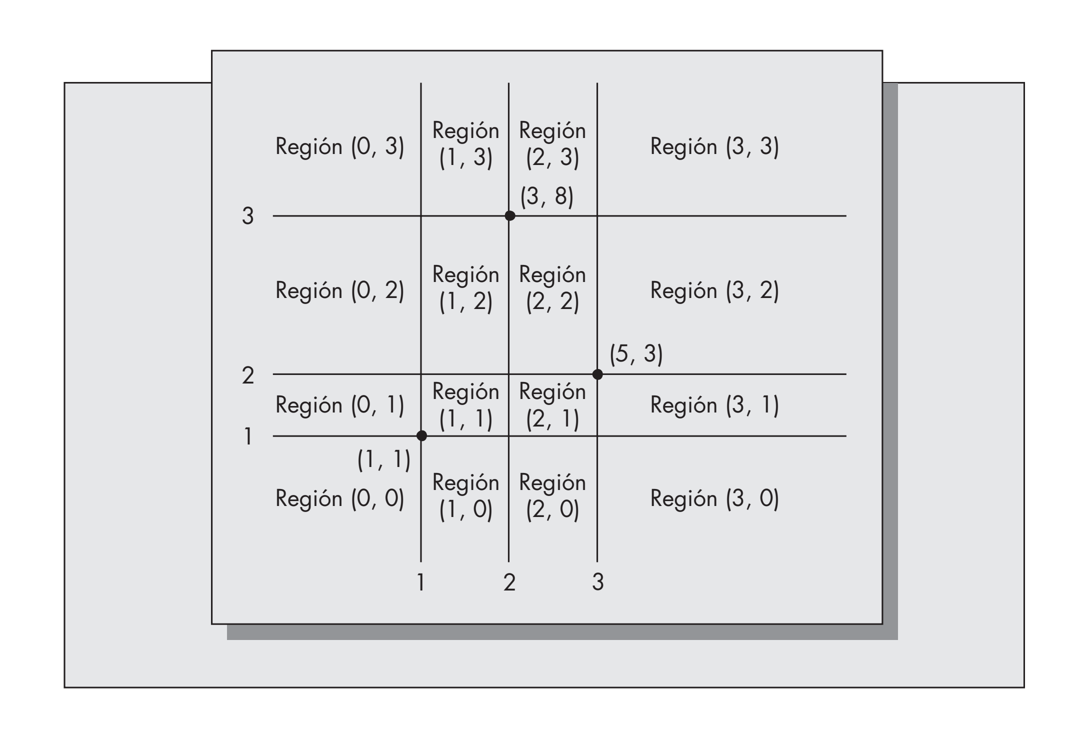


FIGURA 10-37
Regiones para el ejemplo 10B.1.

1, 2 y 3 y tres líneas horizontales etiquetadas también como 1, 2, 3. Las regiones se etiquetan desde (0, 0) hasta (3, 3) como en la figura.

Enseguida calculamos $M_0, \ldots, M_3$ y $N_0, \ldots, N_3$.

$$M_0 = -(2 + 3 + 6) = -11 = N_0,$$

$$M_1 = -11 + (2)(2) = -7, \quad N_1 = -11 + 2(2) = -7,$$

$$M_2 = -7 + (2)(6) = +5, \quad N_2 = -7 + (2)(3) = -1,$$

$$M_3 = +5 + (2)(3) = +11, \quad N_3 = -1 + (2)(6) = +11.$$

A continuación calculamos los cocientes para encontrar la pendiente de cada región.

$$S_{0,0} = -(-11)/(-11) = -1, \qquad S_{0,2} = -(-11)/(-1) = -11,$$

$$S_{1,0} = -(-7)/(-11) = -0.64, \qquad S_{1,2} = -(-7)/(-1) = -7,$$

$$S_{2,0} = -(5)/(-11) = +0.45, \qquad S_{2,2} = -(5)/(-1) = 5,$$

$$S_{3,0} = -(11)/(-11) = +1, \qquad S_{3,2} = -(11)/(-1) = 11,$$

$$S_{0,1} = -(-11)/(-7) = -1.57, \qquad S_{0,3} = -(-11)/(11) = 1,$$

$$S_{1,1} = -(-7)/(-7) = -1, \qquad S_{1,3} = -(-7)/(11) = 0.64,$$

$$S_{2,1} = -(5)/(-7) = +0.71, \qquad S_{2,3} = -(5)/(11) = -0.45,$$

$$S_{3,1} = -(11)/(-7) = +1.57, \qquad S_{3,3} = -(11)/(11) = -1.$$

Antes de construir las curvas de nivel, conviene colocar las pendientes en las regiones apropiadas, como se muestra en la figura 10–38. Una curva de nivel puede iniciarse en cualquier punto del límite de una región. A partir del punto inicial, se dibuja una línea con la pendiente apropiada para esa región hasta el límite de la siguiente región. En ese punto la pendiente cambia al valor asociado con la siguiente región. Se continúa hasta que los segmentos de línea regresen al punto original. (Si las pendientes son correctas y el dibujo es exacto, siempre se regresará al punto de origen.) En la figura 10-39 se muestran dos curvas de nivel típicas para el problema de ejemplo.

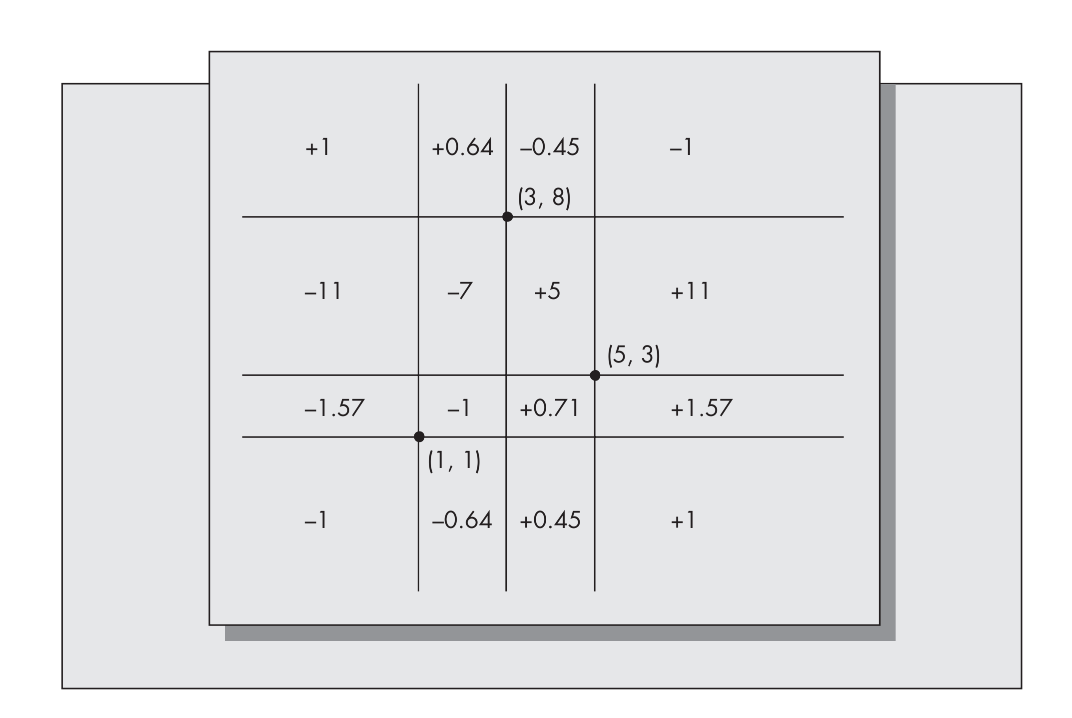
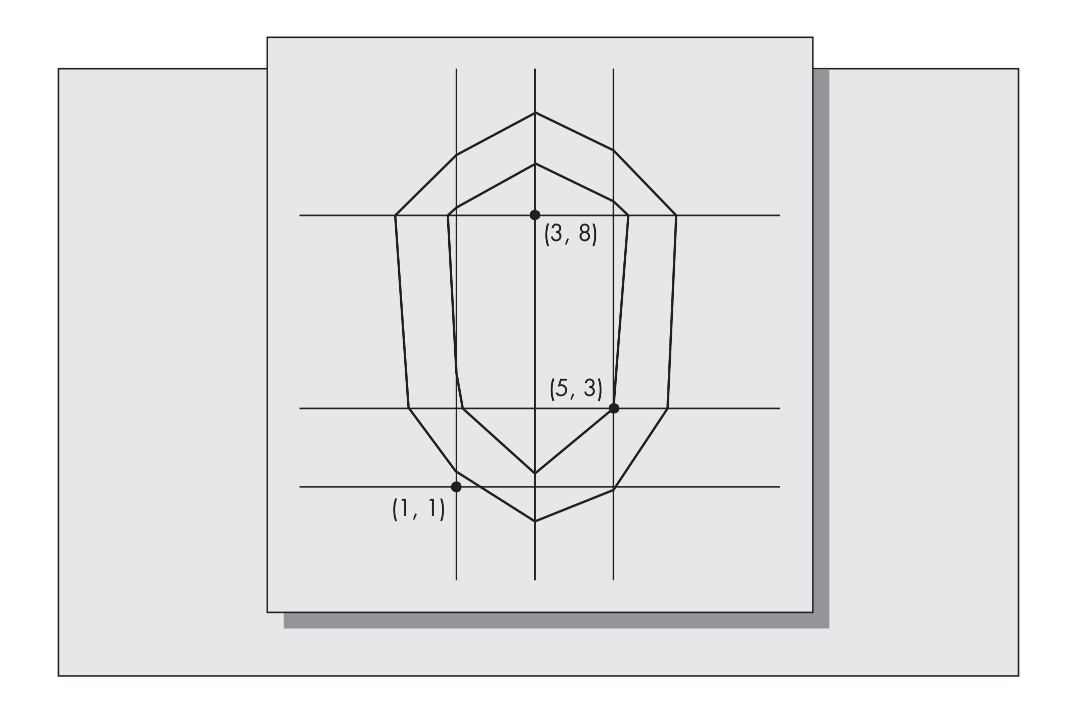


FIGURA 10-38
Pendientes para el ejemplo 10B.1.

FIGURA 10-39
Curvas de nivel para el ejemplo 10B.1.



Bibliografía

Apple, J. M. *Plant Layout and Material Handling*. 3a. ed. Nueva York: John Wiley & Sons, 1977.

Armour, G. C. y E. S. Buffa. "A Heuristic Algorithm and Simulation Approach to Relative Allocation of Facilities". *Management Science* 9 (1963), pp. 294-309.

Arntzen, B. C., G. G. Brown, T. P. Harrison y L. L. Trafton. "Global Supply Chain Management at Digital Equipment Corporation". *Interfaces* 25, núm. 1 (1995), pp. 69-93.

Block, T. E. "A Note on 'Comparison of Computer Algorithms and Visual Based Methods for Plant Layout' por M. Scriabin y R. C. Vergin". *Management Science* 24 (1977), pp. 235-237.

Bruno, G. y P. Biglia. "Performance Evaluation and Validation of Tool Handling in Flexible Manufacturing Systems Using Petri Nets". *Proceedings of the International Workshop on Timed Petri Nets,* pp. 64-71. Turín, Italia, 1985.

Buffa, E. S. "On a Paper by Scriabin and Vergin". *Management Science* 23 (1976), p. 104.

Buffa, E. S., G. C. Armour y T. E. Vollmann. "Allocating Facilities with CRAFT". *Harvard Business Review* 42 (1964), pp. 136-158.

Burbidge, J. L. *The Introduction of Group Technology.* Londres: Heinemann, 1975.

Buzacott, J. A. y D. D. Yao. "On Queuing Network Models of Flexible Manufacturing Systems". *Queuing Systems* 1 (1986), pp. 5-27.

Cohen, M. y H. L. Lee. "Resource Deployment Analysis of Global Manufacturing and Distribution Networks". *Journal of Manufacturing and Operations Management* 2 (1989), pp. 81-104.

Coleman, D. R. "Plant Layout: Computers versus Humans". *Management Science* 24 (1977), pp. 107-112.

Conway, R. W., W. L. Maxwell, J. O. McClain y S. L. Worona. *Users Guide to XCELL + Factory Modeling System.* Redwood City, CA: Scientific Press, 1987.

Cook, N. H. "Computer Managed Parts Manufacture". *Scientific American* 232 (1975), pp. 23-29.

Deisenroth, M. P. y J. M. Apple. "A Computerized Plant Layout Analysis and Evaluation Technique (PLANET)". *Technical Papers 1962.* Norcross, GA: American Institute of Industrial Engineers, 1972.

Drezner, Z. "DISCON: A New Method for the Layout Problem". O*perations Research* 28 (1980), pp. 1375-1384.

Eilon, S., C. D. T. Watson-Gandy y N. Christofides. *Distribution Management: Mathematical Modelling and Practical Analysis.* Londres: Griffin Press, 1971.

Foulds, L. R. "Techniques for Facilities Layout". *Management Science* 29 (1983), pp. 1414-1426.

Francis, R. L., L. F. McGinnis y J. A. White. "Locational Analysis". *European Journal of Operational Research* 12 (1983), pp. 220-252.

Francis, R. L. y J. A. White. *Facility Layout and Location: An Analytical Approach.* Englewood Cliffs, NJ: Prentice Hall, 1974.

Groover, M. P. y E. W. Zimmers. *CAD/CAM: Computer Aided Design and Manufacturing*. Englewood Cliffs, NJ: Prentice Hall, 1984.

Hakimi, S. L. "Optimum Location of Switching Centers and the Absolute Centers and Medians of a Graph". *Operations Research* 12 (1964), pp. 450-459.

Hartley, J. *FMS at Work.* Amsterdam: IFS y North Holland, 1984.

Jaikumar, R. "Flexible Manufacturing Systems: A Management Perspective". Manuscrito inédito, Harvard School of Business, 1984.

Jaikumar, R. "Postindustrial Manufacturing". *Harvard Business Review* 64 (1986), pp. 69-76.

Johnson, Roger V. "SPACECRAFT for Multi-Floor Layout Planning". *Management Science* 28 (1982), pp. 407-417.

Kaplan, R. S. "Must CIM Be Justified by Faith Alone?" *Harvard Business Review* 64 (1986), pp. 69-76.

Keefer, K. B. "Easy Way to Determine the Center of Distribution". *Food Industries* 6 (1934), pp. 450-451.

Krouse, J. "Flexible Manufacturing Systems Begin to Take Hold". *High Technology* 6 (1986), p. 26.

Lee, R. C. y J. M. Moore. "CORELAP-Computerized Relationship Layout Planning". *Journal of Industrial Engineering* 18 (1967), pp. 194-200.

Muther, R. *Practical Plant Layout.* Nueva York: McGraw-Hill, 1955.

Nicol, L. M. y R. H. Hollier. "Plant Layout in Practice". *Material Flow* 1, núm. 3 (1983), pp. 177-188.

Nugent, C. E., T. E. Vollmann y J. Ruml. "An Experimental Comparison of Techniques for the Assignment of Facilities to Locations". *Operations Research* 16 (1968), pp. 150-173.

Pritsker, A. A. B. *Introduction to Simulation and SLAM II.* 3a. ed. Nueva York: John Wiley & Sons, 1986.

Rosenblatt, M. J. "The Dynamics of Plant Layout". *Management Science* 32, núm. 1 (1986), pp. 76-86.

Sassani, F. "A Simulation Study on Performance Improvement of Group Technology Cells". *International Journal of Production Research* 28 (1990), pp. 293-300.

Schweitzer, P. J. "Maximum Throughput in Finite Capacity Open Queuing Networks with Product-Form Solutions". *Management Science* 24 (1977), pp. 217-223.

Scriabin, M. y R. C. Vergin. "Comparison of Computer Algorithms and Visual Based Methods for Plant Layout". *Management Science* 22 (1975), pp. 172-81.

Scriabin, M. y R. C. Vergin. “A Cluster Analytic Approach to Facility Layout”. *Management Science* 31 (1985), pp. 33-49.

Seehof, J. M. y W. O. Evans. “Automated Layout Design Programs”. *Journal of Industrial Engineering* 18 (1967), pp. 690-695.

Sengupta, S. y R. Combes. “Optimizing and General Food’s Environment”. *IIE Solutions,* agosto de 1995, pp. 30-35.

Sly, D. “Computerized Facilities Design and Management: An Overview”. *IIE Solutions,* agosto de 1995, pp. 43-51.

Stecke, K. E. “Formulation and Solution of Nonlinear Integer Production Planning Problems for Flexible Manufacturing Systems”. *Management Science* 29 (1983), pp. 273-288.

Suri, R. y R. R. Hildebrant. “Modeling Flexible Manufacturing Systems Using Mean Value Analysis”. *SME Journal of Manufacturing Systems* 3 (1984), pp. 27-38.

Tansel, B. C., R. L. Francis y T. J. Lowe. “Location on Networks: A Survey (Parts 1 and 2)”. *Management Science* 29 (1983), pp. 482-511.

Tompkins, J. A. y R. Reed Jr. “An Applied Model for the Facilities Design Problem”. *International Journal of Production Research* 14 (1976), pp. 583-595.

Tompkins, J. A. y J. A. White. *Facilities Planning.* Nueva York: John Wiley & Sons, 1984.

Vollman, T. E. y E. S. Buffa. “The Facilities Layout Problem in Perspective”. *Management Science* 12 (1966), pp. 450-468.

Zygmont, J. “Flexible Manufacturing Systems: Curing the Cure-all”. *High Technology* 6 (1986), pp. 22-27.

Capítulo **Once**

Aseguramiento de la calidad

Panorama general del capítulo

Objetivo

Entender el significado de la calidad en el contexto de las operaciones, cómo puede medirse, y cómo puede mejorarse.

Puntos clave

1. *¿Qué es la calidad?* Aun cuando todos tenemos una idea de lo que entendemos por calidad, su definición precisa como una cantidad medible no es fácil. Una definición útil es el cumplimiento de las especificaciones. Esto puede medirse y cuantificarse. Si puede cuantificarse, puede mejorarse. Sin embargo, tal definición no llega a abarcar todos los aspectos de lo que entendemos por calidad, y cómo la percibe el cliente.
2. *Control estadístico de procesos.* Los métodos estadísticos pueden ayudar a monitorear la calidad en el contexto de la manufactura. La base subyacente de las *gráficas de control estadístico* es la distribución normal. La distribución normal (la distribución con forma de campana) tiene la propiedad de que la media más y menos dos desviaciones estándar ($\mu \pm 2\sigma$) contiene aproximadamente el 95 por ciento de la población, y la media más y menos tres desviaciones estándar ($\mu \pm 3\sigma$) contiene más del 99 por ciento de la población. Son estas propiedades las que forman la base de las gráficas de control estadístico. Considere un proceso de manufactura que produce un artículo con una cantidad medible que debe conformar una especificación dada. Se promedian las mediciones de esta cantidad en subgrupos (comúnmente de tamaño cuatro o cinco). El teorema del límite central garantiza que la distribución de la medición promedio tendrá una distribución aproximadamente normal. Si el promedio de un subgrupo se sitúa fuera por dos o tres límites de sigma de la distribución normal, es improbable que tal desviación se deba al azar. Ello apunta a una situación fuera de control, que podría requerir una intervención en el proceso. Esta es la base de la gráfica $\overline{X}$.

   Aun cuando la gráfica $\overline{X}$ es una valiosa manera de probar si hay una desviación en la media subyacente de un proceso, no señala desviaciones en la variación del proceso. Para monitorear la variación del proceso, se calcula el rango de las mediciones de subgrupo (es decir, el mayor valor menos el menor valor en el subgrupo). Como el rango de una muestra es proporcional a la desviación estándar de la muestra, dicho estadístico puede usarse para controlar la variación del proceso. Tal es el propósito de la gráfica *R*, la cual establece los límites de control superior e inferior sobre el rango promedio de los subgrupos y señala cuándo es que la variación del proceso se ha salido de control.

3. *Las gráficas* p *y* c. Las gráficas $\overline{X}$ y *R* son útiles para medir la calidad a lo largo de una dimensión escalar individual como la longitud o el peso. En otros casos, podría haber interés en si el artículo funciona o no. Bajo estas circunstancias, la gráfica *p* es la apropiada. Ésta se basa en la distribución binomial. Un artículo tiene el atributo apropiado o no lo tiene. Si el valor observado de *p* (la proporción de artículos en buen estado) experimenta una desviación repentina, éste señala una posible situación de pérdida del control.

   La gráfica *c* se basa en la distribución de Poisson. La distribución de Poisson describe eventos que ocurren completamente en forma aleatoria con respecto al tiempo o al espacio. En el contexto del control de calidad estadístico, considere una situación en la cual es aceptable un cierto número de defectos, como las abolladuras menores en un automóvil, entre los cuales hay muchos que se consideran inaceptables. En este caso, la gráfica *c* sería un medio apropiado para controlar el proceso. El parámetro *c* es la tasa promedio de ocurrencia de defectos, y una señal de pérdida del control se activa si el valor observado de c es demasiado alto. Observe que comúnmente las dos gráficas *p* y *c* se implementan con una distribución normal ya que, bajo las circunstancias adecuadas, la distribución normal provee una buena aproximación a las dos distribuciones, binomial y de Poisson.

4. *Diseño económico de las gráficas de control.* El control de calidad estadístico requiere varios pasos, cada uno de los cuales incurre en un costo diferente. Primero está el costo de inspeccionar los artículos. Para las gráficas $\overline{X}$, suponemos muestras de subgrupo de tamaño *n*. Así, cada muestreo de subgrupo incurre en un costo proporcional a *n*. Segundo, si se detecta una situación de pérdida de control, el costo de investigar la causa del problema puede ser elevado. Aunque la señal de pérdida de control sea una falsa alarma, el proceso debe detenerse. Finalmente, si el proceso continúa operando en un estado de pérdida de control, esto también podría conducir a costos elevados a medida que aumentan los inventarios de artículos defectuosos. Pueden escogerse límites de control que equilibren estos costos de la mejor manera.

5. *Muestreo de aceptación.* La segunda parte de este capítulo trata sobre el muestreo de aceptación. El *muestreo de aceptación* ocurre después que se ha producido un lote de artículos, más bien que durante el proceso de manufactura. Puede ser realizado por el fabricante o por el consumidor. En la mayoría de los casos, es impráctico, imposible o demasiado costoso inspeccionar al 100 por ciento los artículos. Por estas razones, un enfoque más común es muestrear un subconjunto del lote y escoger la aceptación o el rechazo del lote con base en los resultados del muestreo. Los planes de muestreo más comunes son (1) el muestreo simple, (2) el muestreo doble, y (3) el muestreo secuencial.

   En el caso del muestreo simple, se seleccionan n artículos de un lote de *N* artículos (donde $n < N$) y se rechaza el lote si el número de los defectos sobrepasa cierto nivel. El muestreo doble implica que si el número de artículos defectuosos se sitúa entre dos límites previamente especificados (es decir, no es ni muy alto ni muy bajo), se hace un nuevo muestreo para determinar el destino del lote. En el muestreo secuencial se decide ya sea aceptar el lote, rechazar el lote, o continuar el muestreo después que cada artículo es muestreado. Los límites apropiados para cada una de estas pruebas se basan en las distribuciones subyacentes de probabilidad y en la especificación de niveles aceptables del error Tipo 1 (α).

6. *Administración de la calidad total.* A medida que el movimiento de la calidad comenzó a arraigarse en Estados Unidos y en otras partes del mundo, una manera en que se describía el compromiso de una organización con la calidad era la administración de la calidad total (TQM, por las siglas en inglés de *total quality management*). Brevemente, éste es el compromiso completo de todas las partes de una compañía con la misión de la calidad. Una parte importante de TQM reside en escuchar al consumidor. Este proceso incluye las encuestas a clientes y grupos de enfoque para investigar qué es lo que quiere

el cliente, destilar esta información, priorizar las necesidades del cliente, y vincular estas necesidades con el diseño del producto. Una manera de alcanzar el último término de la lista es satisfacer la función de calidad (QFD, por sus siglas en inglés).

Varias agencias en todo el mundo promueven la calidad en sus respectivos países a través del reconocimiento formal. Este proceso comenzó en Japón con el premio Deming, establecido y fundado por el gurú de la calidad, W. Edwards Deming. En Estados Unidos, se reconoce la calidad extraordinaria con el Premio Baldrige. Otro desarrollo importante es la certificación de la Organización Internacional de Estándares, ISO 9000, que requiere que las compañías documenten claramente sus políticas y procedimientos. Aun cuando el proceso de certificación puede ser costoso tanto en tiempo como en dinero, con frecuencia es necesario para hacer negocios en muchos países.

El capítulo concluye con un análisis del diseño para la calidad. Al realizar una mayor inversión frente a un diseño sólido de producto, el consumidor será recompensado con productos superiores y la compañía será retribuida con la lealtad del cliente.

Aun cuando la economía estadounidense fue sólida durante la segunda mitad de la década de los noventa, existen algunas tendencias perturbadoras. Su balanza comercial en bienes manufacturados sigue siendo negativa año tras año. En especial, han conservado una balanza comercial negativa con Japón durante varias décadas, aun durante la recesión que asoló a ese país en la década de los noventa. Esta balanza comercial negativa implica que Estados Unidos importa más de lo que exporta. En el caso de Japón, esto se debe en gran parte al hecho de que los estadounidenses están consumiendo productos hechos en Japón a una tasa creciente. El éxito de Japón en los productos electrónicos para el consumidor y en los automóviles son solamente dos ejemplos. Pero no es sólo el consumo de productos hechos en Japón lo que está aumentando. Otros ejemplos son los automóviles hechos en Alemania y en Suecia. Los artículos para la cocina hechos en Alemania, las bicicletas hechas en Taiwán, y los relojes de lujo hechos en Suiza. En muchos casos, los productos importados son considerablemente más caros que sus contrapartes estadounidenses, y aún así tienen preferencia. ¿Por qué? La respuesta simple es que se percibe que tienen mayor calidad.

¿Qué es la calidad? El pensamiento tradicional nos diría que la calidad es cumplir con las especificaciones; es decir, ¿hace el producto aquello para lo que fue diseñado? Algunos piensan que esta definición es la única definición significativa de la calidad, puesto que el cumplimiento es algo que puede medirse. De acuerdo con Philip Crosby (1979),

> Esa es precisamente la razón por la cual debemos definir a la calidad como "el apego a los requerimientos" si vamos a administrarla. Sí, aquellos que quieran hablar sobre la calidad de la vida deben hablar acerca de esa vida en términos específicos, como ingreso deseable, salud, control de la contaminación, programas políticos, y otros aspectos que puedan ser medidos.

Crosby señala un buen punto. Al definir la calidad en términos del apego, evitamos hacer comparaciones poco razonables. ¿Es un Rolls-Royce un producto de mejor calidad que un Toyota Corolla? No necesariamente. El Toyota puede ser un producto de mejor calidad en relación con *lo que fue diseñado para hacer*.

Sin embargo, esto no nos cuenta la historia completa. Así como la belleza está en los ojos de quien contempla, la calidad está en la mente del cliente. Si el consumidor no está feliz con el producto, no es de alta calidad. Vista de esta manera, la calidad es una medida del apego del producto a las *necesidades del cliente*.

¿Por qué esto es diferente? El apego a las especificaciones supone un diseño dado y las especificaciones resultantes de ese diseño. La satisfacción de las necesidades del cliente implica que el diseño del producto es parte de la evaluación. Dadas dos máquinas lavadoras con necesidades comparables de mantenimiento, ¿qué determina cuál es la que va a comprar el consumidor? La respuesta es una combinación de estética, características y diseño. Una visión

amplia de la calidad como ésta es tanto buena como mala. Es buena porque llega al núcleo del problema: la calidad es lo que el cliente piensa que es. Es mala en que hace difícil medir la calidad y por tanto es difícil mejorarla.

Casi no hay duda de que la confiabilidad del producto es una parte importante del espectacular éxito de los fabricantes de automóviles japoneses. Aunque varios fabricantes de automóviles estadounidenses gozan de tasas iniciales de defectos comparables a las de los automóviles japoneses (como lo mide la encuesta de compradores de automóviles nuevos J. D. Power), no es el número de defectos en los automóviles nuevos lo que realmente importa para la mayoría de los clientes. Es la confiabilidad del automóvil a lo largo de su vida. En este aspecto, pocos fabricantes de automóviles pueden acercarse a las puntuaciones logradas por Honda y Toyota y otros fabricantes de automóviles japoneses. Cada año la Unión de Consumidores realiza una encuesta de su presteza para determinar las experiencias de los lectores ese año con los productos que tienen. La confiabilidad de los automóviles se mide de acuerdo con una escala de cinco puntos: mucho peor que el promedio, peor que el promedio, promedio, mejor que el promedio, y mucho mejor que el promedio. Las graduaciones corresponden a varios factores, incluyendo fallas mecánicas, fallas eléctricas, e integridad de la carrocería. Los automóviles japoneses han alcanzado puntuaciones consistentemente mejores que la mayoría de sus competidores estadounidenses y europeos.

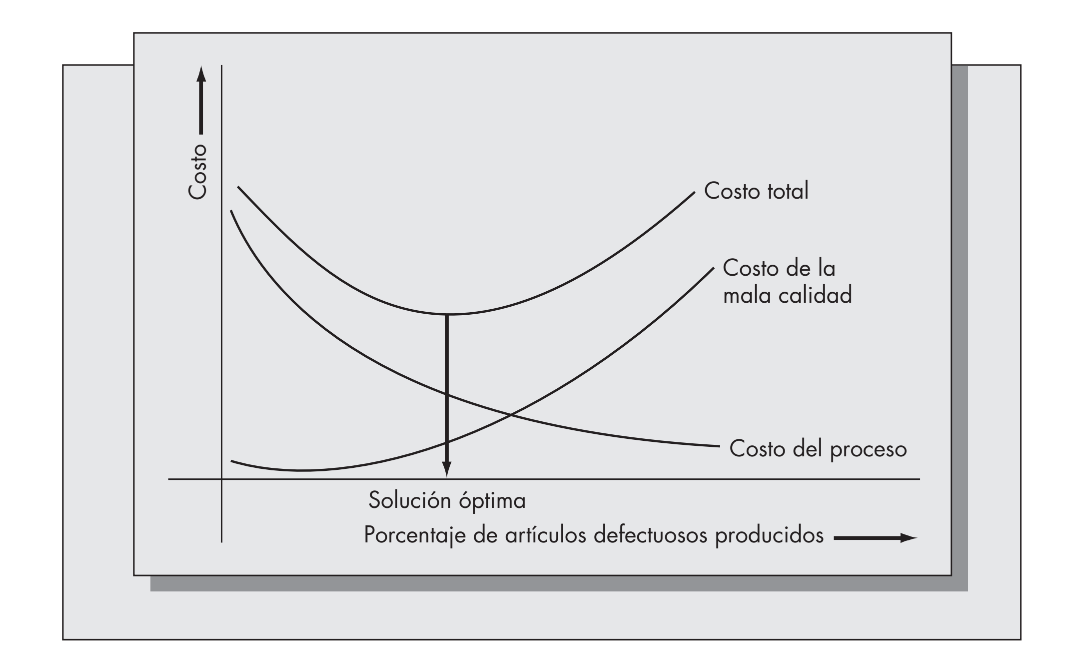
En la década de los años cincuenta virtualmente todos los automóviles vendidos en Estados Unidos eran de manufactura estadounidense. En 1955 tres fabricantes grandes de automóviles (Ford, GM y Chrysler) representaban 95 por ciento de las ventas en Estados Unidos y la mayor parte del 5 por ciento restante eran ventas hechas por otras marcas estadounidenses (ahora extintas). Actualmente, las compañías japonesas representan cerca de 30 por ciento de las ventas en Estados Unidos y están haciendo aún más progresos en ventas a nivel mundial. En 1961 los fabricantes de automóviles estadounidenses representaban casi 50 por ciento del mercado mundial de vehículos de pasajeros, y Japón aproximadamente 2 por ciento. En la actualidad, los fabricantes estadounidenses representan aproximadamente 15 por ciento del mercado mundial de automóviles de pasajeros, mientras que la participación de Japón en el mercado supera 20 por ciento. (Sin embargo, Estados Unidos ha seguido incrementando su participación en el mercado de vehículos comerciales, representando a la fecha aproximadamente 40 por ciento del total mundial.)

Toyota en especial sigue haciendo incursiones importantes en el mercado estadounidense. La tercera posición de Chrysler en Estados Unidos está siendo amenazada por Toyota, y es evidente que si las tendencias presentes continúan, Toyota va a superar dentro de poco a Chrysler. De hecho, cuando se escribió este libro, a finales de 2003, el Toyota Camry había sido el modelo vendedor número uno en cinco de los pasados seis años (con el Honda Accord en segundo lugar). Además, tanto Toyota como Honda han sobresalido como líderes de tecnología con los primeros híbridos de gas-electricidad disponibles en Estados Unidos. Aun cuando se pueda pensar que los híbridos son una parte insignificante del mercado, los distribuidores tuvieron 10 000 órdenes del Toyota Prius 2004 rediseñado antes de que los automóviles estuvieran disponibles en el otoño de 2003 (*Business Week*, 17 de noviembre de 2003). Aún más, Lexus, la división de lujo de Toyota, está planeando anunciar tanto un híbrido de lujo como una SUV híbrida en los años venideros. [Los híbridos actuales ofrecen más de 50 millas por galón y se clasifican como vehículos súper ultra de bajas emisiones (SULEV).]

Los progresos enormes logrados por los japoneses en electrónica y en automóviles para los consumidores dan la impresión de que los productos estadounidenses y la industria estadounidense son inferiores. Esto simplemente no es verdad. Las compañías de origen estadounidense son líderes en muchas industrias importantes. La construcción de aviones, las computadoras centrales, la biotecnología, los servicios financieros, los electrodomésticos, los productos químicos, y las telecomunicaciones son una pequeña muestra de las industrias que dominan las compañías estadounidenses. Transfiriendo lecciones aprendidas en estas industrias, los estadounidenses comenzar a recobrar su margen competitivo en la manufactura en general.

La administración debe encarar problemas difíciles como saber cuánto invertir en la calidad y determinar la mejor manera de proceder al hacer esa inversión.

FIGURA 11-1
El compromiso entre calidad y costo.



Existe una combinación óptima entre el costo de la mala calidad y la inversión requerida en el proceso para mejorar la calidad, como se representa en la figura 11-1. Aunque en principio estas curvas pueden dibujarse, es difícil evaluar los costos de la mala calidad. Los costos directos, tales como los que resultan del desperdicio, el reprocesamiento y la inspección, son relativamente fáciles de precisar. Pero, ¿cómo determinar lo que cuesta perder la lealtad del consumidor? Nadie negaría que la publicidad es esencial, pero ¿el énfasis que se ha dado a la mercadotecnia en Estados Unidos ha afectado la manufactura? Debemos reconocer que el consumidor moderno está más educado y discrimina más que nunca antes. Los anuncios ingeniosos no venderán productos de segunda clase. Es tiempo de colocar nuestra inversión donde pertenece: en el diseño y la manufactura de productos de calidad.

Panorama general del capítulo

El control estadístico de la calidad se remonta a la década de los años treinta. Sus raíces radican en el trabajo de Walter Shewhart, un científico empleado en los laboratorios Bell Telephone. W. Edwards Deming, el hombre a quien se le atribuye haber llevado el mensaje del control de la calidad a los japoneses, era un estudiante de Shewhart. Deming ha enfatizado en sus enseñanzas que comprender el concepto de la variación estadística de los procesos es un paso clave para el diseño de un programa efectivo de control de la calidad. Es necesario entender la variación del proceso con objeto de saber cómo producir productos conforme a las especificaciones. Deming se ha convertido en un semidiós en Japón, donde sus enseñanzas suscitaron la revolución japonesa de la calidad.

Este capítulo está dirigido a suministrar a los estudiantes una comprensión de los aspectos esenciales que subyacen al control estadístico de la calidad. Estudiamos dos áreas básicas: las *gráficas de control* y el *muestreo de aceptación*. Brevemente, una gráfica de control es un medio gráfico para determinar si la distribución subyacente de alguna variable medible pareciera haber experimentado una desviación. El muestreo de aceptación es el conjunto de procedimientos para inferir las características de un lote a partir de las características de una muestra de artículos de ese lote. Aun cuando el pensamiento de "cero defectos" podría sugerir que estos enfoques son obsoletos, nosotros diferimos. Continúan en uso el monitoreo de procesos y el muestreo estadístico. Es importante entender la teoría subyacente detrás de estos métodos para saber cuándo y cómo deben aplicarse.

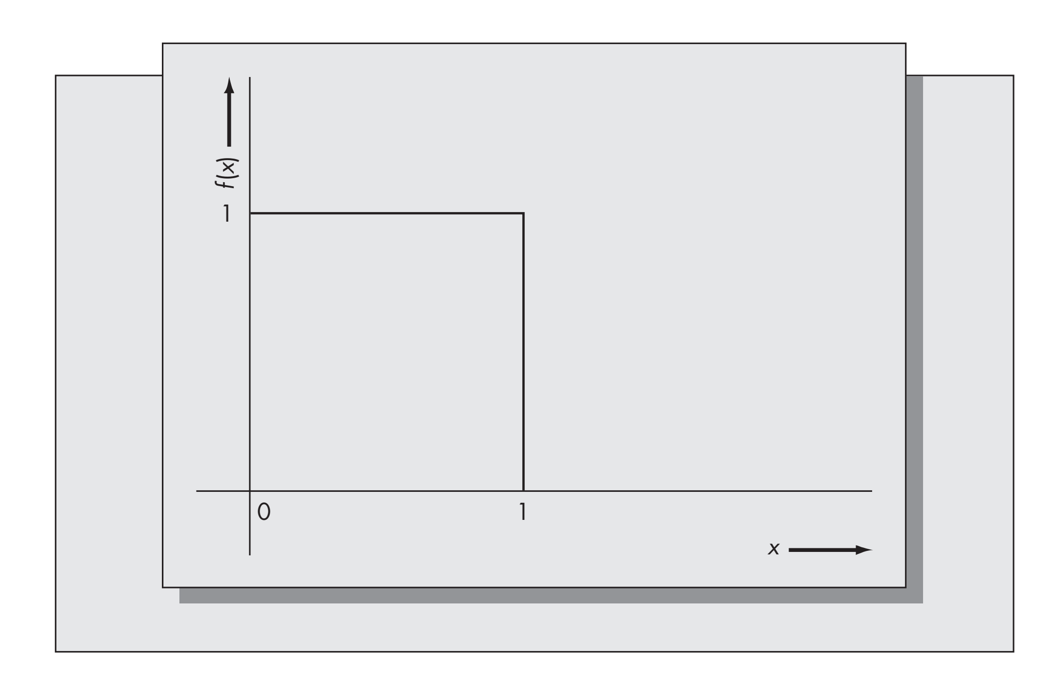
11.1 BASES ESTADÍSTICAS DE LAS GRÁFICAS DE CONTROL

Las *gráficas de control* proporcionan un medio gráfico simple para controlar un proceso en tiempo real. Aunque son fáciles de construir y fáciles de usar, las gráficas de control se basan en principios estadísticos rigurosos. Han ganado amplia aceptación en la industria y se prefieren a métodos estadísticos más convencionales.

Una gráfica de control describe la salida de un proceso de producción con respecto al tiempo y señala cuándo pudo haber ocurrido un cambio en la distribución de probabilidad que genera las observaciones. Para construir una gráfica de control se usa información sobre la distribución de probabilidad de la variación del proceso y los resultados fundamentales de la teoría de probabilidad. Un resultado que constituye la base de una clase de gráficas de control se conoce como el *teorema del límite central*. A grandes rasgos, el teorema del límite central dice que la distribución de las sumas de las variables aleatorias independientes y con distribución idéntica tiende a la distribución normal a medida que aumenta el número de términos de la suma.[1] En general, la distribución de la suma converge muy rápidamente con una distribución normal. Para ilustrar el teorema del límite central, suponga que X es una variable aleatoria con la distribución uniforme en el intervalo (0, 1). La función de densidad de probabilidad de X se ilustra en la figura 11-2.

La densidad de X tiene poca semejanza con una densidad normal. Ahora supongamos que las tres variables aleatorias X_1, X_2 y X_3 son variables aleatorias independientes, cada una de las cuales tiene la distribución uniforme en el intervalo (0, 1). Considere la variable aleatoria $W = X_1 + X_2 + X_3$. La distribución de W puede deducirse mediante la convolución de las distribuciones de X_1, X_2, y X_3. Los detalles no se presentarán aquí. (El lector interesado puede consultar un libro a nivel de postgrado sobre probabilidad como DeGroot, 1986.) La función de densidad de probabilidades de W se muestra en la figura 11-3. Ahora se nota bastante la similitud con una densidad normal. En la figura 11-4 hemos graficado la función de densidad de probabilidades de W junto con la aproximación normal asociada. Observe cómo concuerdan las dos curvas. Si siguiéramos agregando variables aleatorias uniformes independientes, la concordancia sería aún mayor.

FIGURA 11-2
Densidad de probabilidades de una variable uniforme en (0, 1).



[1] El teorema del límite central también se usó en el capítulo 9 para justificar el uso de la distribución normal para describir el tiempo de terminación de un proyecto en las redes PERT.

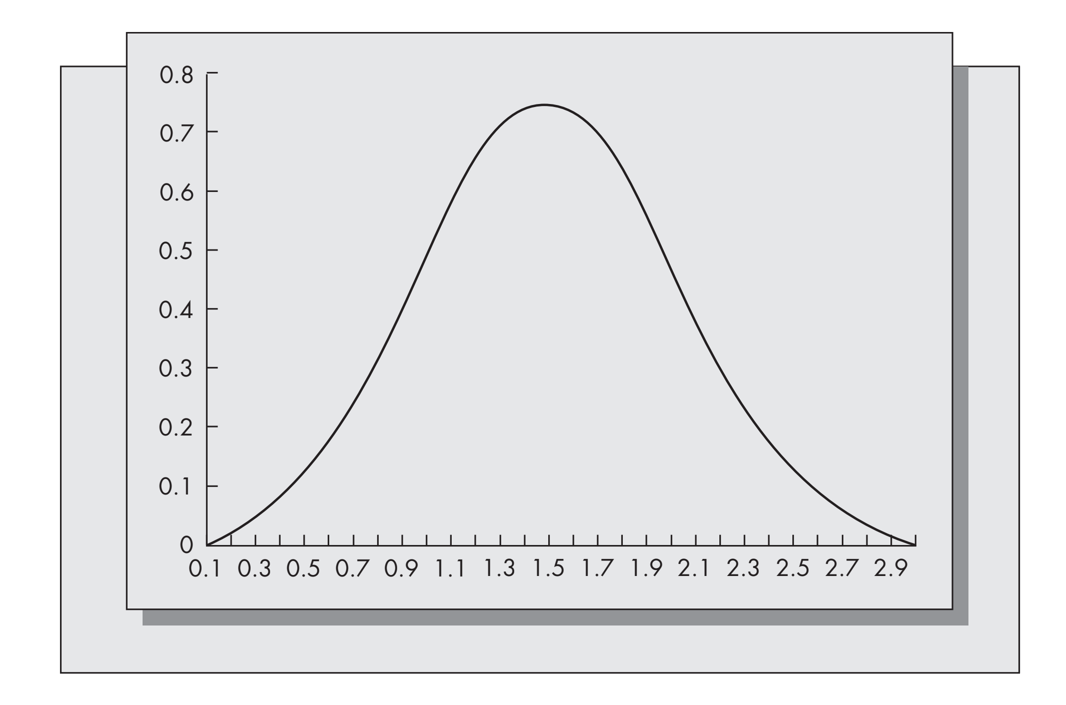


FIGURA 11-3
Densidad de la suma de tres variables aleatorias uniformes.

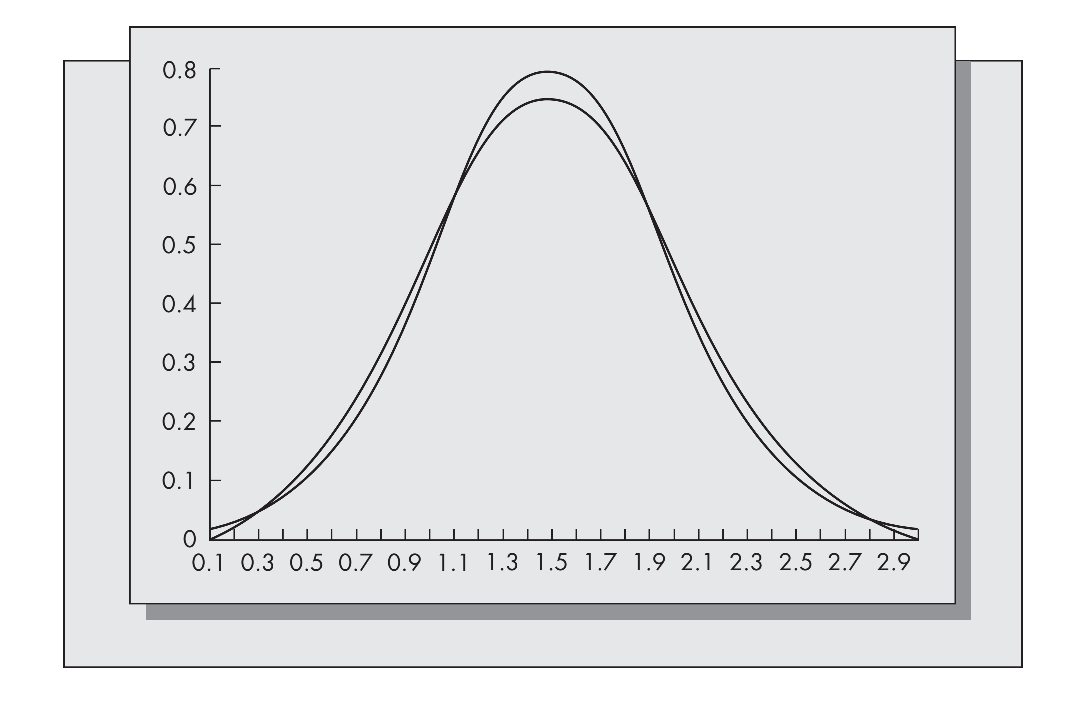


FIGURA 11-4
Densidad de la suma de tres variables aleatorias uniformes y la aproximación normal.

En el control de la calidad, el teorema del límite central justifica la hipótesis de que la distribución de $\overline{X}$, la media de la muestra, tiene una distribución aproximadamente normal. Recuerde la definición de la media de la muestra: si $(X_1, X_2, \ldots, X_n)$ es una muestra aleatoria, entonces la media de la muestra $\overline{X}$ se define como

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Suponga que una variable Z tiene la distribución normal estándar. Entonces, a partir de la tabla A-1 al final de este libro,

$$P\{-3 \leq Z \leq 3\} = .9974.$$

En palabras, esto significa que la posibilidad de obtener un valor de Z ya sea mayor que 3 o menor que -3 es .0026, o aproximadamente 3 posibilidades de 1,000. Esta es la base de los así llamados límites tres sigma que se han convertido en el estándar de facto en el control de la calidad. Ahora considere la media de la muestra $\overline{X}$, sobre la cual el teorema del límite central nos dice que tiene una distribución aproximadamente normal. Suponga que la media de cada valor de la muestra es μ y que la desviación estándar de cada valor de la muestra es σ. Entonces es bien sabido que la media de $\overline{X}$ también es μ y la desviación estándar de $\overline{X}$ es $\sigma/\sqrt{n}$. Por lo tanto, la variable estandarizada

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

tiene (aproximadamente) la distribución normal con media cero y varianza unitaria. Se concluye que

$$P\left\{-3 \leq \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq 3\right\} = .9974,$$

que equivale a

$$P\left\{\mu - \frac{3\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + \frac{3\sigma}{\sqrt{n}}\right\} = .9974.$$

Es decir, la posibilidad de observar un valor de $\overline{X}$ mayor que $\mu + 3\sigma/\sqrt{n}$ o menor que $\mu - 3\sigma/\sqrt{n}$ es .0026. Un evento de este tipo es lo suficientemente raro como para que si ocurriera, sería probable que hubiera sido causado por una desviación de la media poblacional, μ, más que haber resultado del azar. Esta es la base de la teoría de las gráficas de control.

Problemas para la sección 11.1

1. Suponga que X_1 y X_2 son variables aleatorias independientes con distribución uniforme en (0, 1, 2, 3, 4, 5). Es decir,

$$f(j) = P\{X_i = j\} = \tfrac{1}{6} \qquad \text{para } i = 1, 2 \text{ y } 0 \leq j \leq 5.$$

a) Calcule $E(X_1)$ y $\text{Var}(X_1)$. [Sugerencia: use las fórmulas

$$E(X) = \sum jf(j),$$
$$\text{Var}(X) = \sum j^2 f(j) - (E(X))^2.]$$

b) Determine la distribución de probabilidad de $Y = X_1 + X_2$. [Sugerencia: para cada valor posible de Y, determine todas las combinaciones de X_1 y X_2 que arrojen ese valor. Por ejemplo, $Y = 3$ puede obtenerse mediante $(X_1, X_2) = (0, 3), (1, 2), (2, 1)$, y $(3, 0)$. Como cada par tiene probabilidad $(\frac{1}{6})(\frac{1}{6}) = \frac{1}{36}$, obtenemos $P\{Y = 3\} = \frac{4}{36} = \frac{1}{9}$. Repita este proceso para todos los valores de Y. Como comprobación asegúrese de que

$$\sum_y P\{Y = y] = 1.0.]$$

c) Usando los resultados del inciso *b*, encuentre $P\{1.5 < Y < 6.5\}$.

d) Usando los resultados $E(Y) = 2E(X_1)$ y $\text{Var}(Y) = 2\text{Var}(X_1)$, aproxime la respuesta para el inciso *c* usando una distribución normal.

e) Suponga que $X_1, X_2, \ldots, X_{20}$ son variables aleatorias independientes con distribución idéntica que tienen distribución uniforme en (0, 1, 2, 3, 4, 5). Usando una aproximación normal, estime

$$P\left\{\sum_{i=1}^{20} X_i \leq 75\right\}.$$

f) ¿Piensa que la aproximación calculada en el inciso *d* o en el inciso *e* sea más exacta? ¿Por qué?

2. Los siguientes datos representan el número de discos defectuosos producidos por hora basándose en la observación del sistema durante 30 horas sucesivas.

0	3	5	2	6	8	3	5	4	6	6	9	5	1	2
1	2	5	3	3	0	1	0	7	1	7	5	4	4	3

a) Grafique un histograma de frecuencias para el número de fallas por hora.

b) Calcule la media y la varianza de la muestra. Use las fórmulas

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right).$$

(Nota: Los cálculos pueden agilizarse agrupando primero los datos, ya que sólo hay 10 valores diferentes.)

c) Si el número de discos defectuosos producidos por hora tiene distribución normal con la media y la varianza calculadas en el inciso *b*, determine la probabilidad de que se observen menos de cinco artículos defectuosos para cualquier hora específica.

d) Si la producción de cada disco le cuesta a la compañía \$5 y se descartan los discos defectuosos, ¿cuánto tiempo tomaría (en el sentido de valor esperado) financiar una nueva pieza de equipo con un costo de \$45 000 que reduciría los artículos defectuosos a la mitad del nivel actual? Suponga semanas de 40 horas de producción y 48 semanas por año.

3. La resistencia a la tensión de las bolsas de plástico de servicio pesado que se usan en los compactadores de basura tiene distribución normal con media de 150 libras por pulgada cuadrada y desviación estándar de 12 libras por pulgada cuadrada. Un contratista independiente las usa para remolcar basura que requiere una resistencia a la tensión de 120 libras por pulgada cuadrada. ¿Qué proporción de las bolsas del compactador no cumplirán con tales requerimientos?

4. Una compañía de evaluación de crédito recomienda otorgar tarjetas de crédito con base en diversos criterios. Uno es el ingreso anual. Si el ingreso anual de los solicitantes tiene distribución normal con media de \$22 000 y desviación estándar de \$4 800 y la compañía no recomienda a ningún solicitante a menos que su ingreso sobrepase los \$15 000, ¿qué fracción de los solicitantes son rechazados sobre esta base?

5. *a*) ¿Cuál es la probabilidad de que una variable normal sobrepase los límites de dos sigma? (Es decir, ¿cuál es la probabilidad de observar un valor de la variable aleatoria mayor que $\mu + 2\sigma$ o menor que $\mu - 2\sigma$?)

b) Si "decidir que el proceso está fuera de control" implica observar el hecho de que la media de la muestra sobrepase límites de *k* sigma, describa las ventajas y desventajas de usar $k = 2$ contra $k = 3$.

6. Los miembros de un club de golf privado tienen handicaps que se distribuyen normalmente con media 15 y desviación estándar 3.5. En un evento específico, escogen grupos de cuatro jugadores elegidos aleatoriamente del club. El handicap del grupo de cuatro es el promedio aritmético de los handicaps de los cuatro jugadores que constituyen el grupo. ¿En que proporción de los grupos será el handicap del grupo de cuatro menor que 10 o mayor que 20? (Sugerencia: la desviación estándar del promedio de cuatro variables aleatorias independientes con distribución idéntica es exactamente la mitad de la desviación estándar de una de ellas.)

11.2 GRÁFICAS DE CONTROL PARA VARIABLES: LAS GRÁFICAS $\overline{X}$ Y *R*

Un proceso está bajo control si opera un sistema estable de causas al azar. Es decir, la distribución de probabilidad básica que genera las observaciones no cambia con el tiempo. Si el valor observado de la media de la muestra de un grupo de observaciones se sitúa fuera de los límites correspondientes de tres sigma, es posible que haya habido un cambio en la distribución de probabilidad que genera las observaciones. Para ilustrar cómo se desarrollan y se interpretan las gráficas de control, considere el siguiente caso de estudio.

Ejemplo *11.1* Wonderdisk produce una línea de discos compatibles que pueden conectarse en equipo de IBM. El edificio 35 es el responsable de la producción de los brazos de lectura/escritura para el disco modelo A55C. Los brazos tienen una longitud aproximada de 2.875 pulgadas. Los ingenieros de diseño han establecido una tolerancia de ± 0.025 pulgadas para las longitudes de los brazos y venden esta cifra en las especificaciones publicadas.

Generalmente la compañía produce 40 brazos por día. En 30 días consecutivos de producción, se muestrean y se miden aleatoriamente cinco brazos de cada día de producción. Las mediciones resultantes se muestran en la tabla 11-1.

Estas observaciones muestran que existe alguna variación en la longitud de los brazos. Sin embargo, no hay ningún patrón discernible para esta variación. Defina a la variable aleatoria X como la longitud de un brazo seleccionado aleatoriamente. Entonces podemos interpretar a la tabla 11-1 como 150 observaciones independientes de la variable aleatoria X.

A Howard Hamilton, un ingeniero industrial que trabaja para Wonderdisk, se le asigna el trabajo de analizar estos datos. Lo primero que observa Howard es que con frecuencia se excedieron las tolerancias establecidas de ± 0.025 pulgadas. Esto se hace muy evidente al calcular el rango de las observaciones diarias. El rango de una muestra es la observación máxima menos la observación mínima. Para los 30 días de datos, Howard observa que el rango sobrepasa a 0.05 en cuatro casos. Para obtener una idea más clara de qué proporción de la población se sitúa fuera de las tolerancias especificadas, Howard desarrolla un histograma de frecuencias de las 150 mediciones. Este histograma se muestra en la figura 11-5. El histograma sugiere que las mediciones tienen distribución normal. Howard usó una prueba de bondad de ajuste para verificar la normalidad de las observaciones. La figura 11-6 muestra la curva teórica normal.

La curva teórica normal se determina de la siguiente manera. Como la distribución normal depende de dos parámetros, μ y σ, debemos estimar estos valores a partir de los datos muestrales. De la teoría de estadística, sabemos que los "mejores" estimados de la media y la varianza de la población son la media de la muestra, $\overline{X}$, y la varianza de la muestra, s^2, dadas por

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \overline{X})^2.$$

TABLA 11-1 Datos para los brazos de lectura/escritura

Número de muestra	Medición de la longitud de un brazo de lectura/escritura					Promedio	Rango
1	2.8971	2.8477	2.8624	2.8606	2.8971	2.8730	.0494
2	2.8863	2.8541	2.8677	2.8838	2.8854	2.8755	.0322
3	2.8772	2.8708	2.8920	2.8892	2.8840	2.8826	.0212
4	2.8808	2.8650	2.8686	2.8874	2.8804	2.8764	.0224
5	2.8633	2.8993	2.8650	2.8909	2.9131	2.8863	.0497
6	2.8743	2.8571	2.8863	2.8473	2.8739	2.8678	.0390
7	2.8820	2.8612	2.8805	2.8737	2.8933	2.8781	.0322
8	2.8847	2.8630	2.8846	2.8969	2.8916	2.8842	.0339
9	2.8569	2.8934	2.8926	2.8585	2.8721	2.8747	.0365
10	2.8784	2.8795	2.8794	2.8608	2.8672	2.8731	.0187
11	2.8821	2.8544	2.9053	2.8495	2.8670	2.8717	.0558
12	2.8643	2.8533	2.8718	2.8565	2.8724	2.8637	.0191
13	2.8675	2.8578	2.8971	2.8709	2.8908	2.8768	.0394
14	2.8495	2.8701	2.8741	2.8699	2.8766	2.8680	.0271
15	2.8822	2.8731	2.8551	2.8782	2.8687	2.8714	.0271
16	2.8731	2.8675	2.8743	2.8520	2.8900	2.8714	.0379
17	2.9054	2.9190	2.8752	2.8477	2.8639	2.8822	.0713
18	2.8759	2.8832	2.8660	2.8667	2.8674	2.8718	.0172
19	2.8676	2.8775	2.8793	2.8943	2.9048	2.8847	.0373
20	2.8765	2.8613	2.8737	2.8524	2.8767	2.8681	.0243
21	2.9052	2.8851	2.8895	2.8904	2.8723	2.8885	.0328
22	2.8606	2.8837	2.9017	2.8628	2.8455	2.8709	.0562
23	2.8752	2.8722	2.8618	2.8637	2.8725	2.8691	.0133
24	2.8566	2.8929	2.9035	2.9109	2.8594	2.8847	.0543
25	2.8495	2.8749	2.8873	2.8557	2.8673	2.8669	.0378
26	2.8736	2.8606	2.8797	2.8522	2.8802	2.8693	.0280
27	2.8449	2.8908	2.8851	2.8798	2.8610	2.8723	.0459
28	2.8589	2.8800	2.9025	2.8974	2.8606	2.8799	.0437
29	2.8910	2.8546	2.8744	2.8775	2.8634	2.8722	.0364
30	2.8607	2.8769	2.8771	2.8934	2.8706	2.8757	.0326

FIGURA 11-5
Histograma de frecuencias de 150 mediciones.

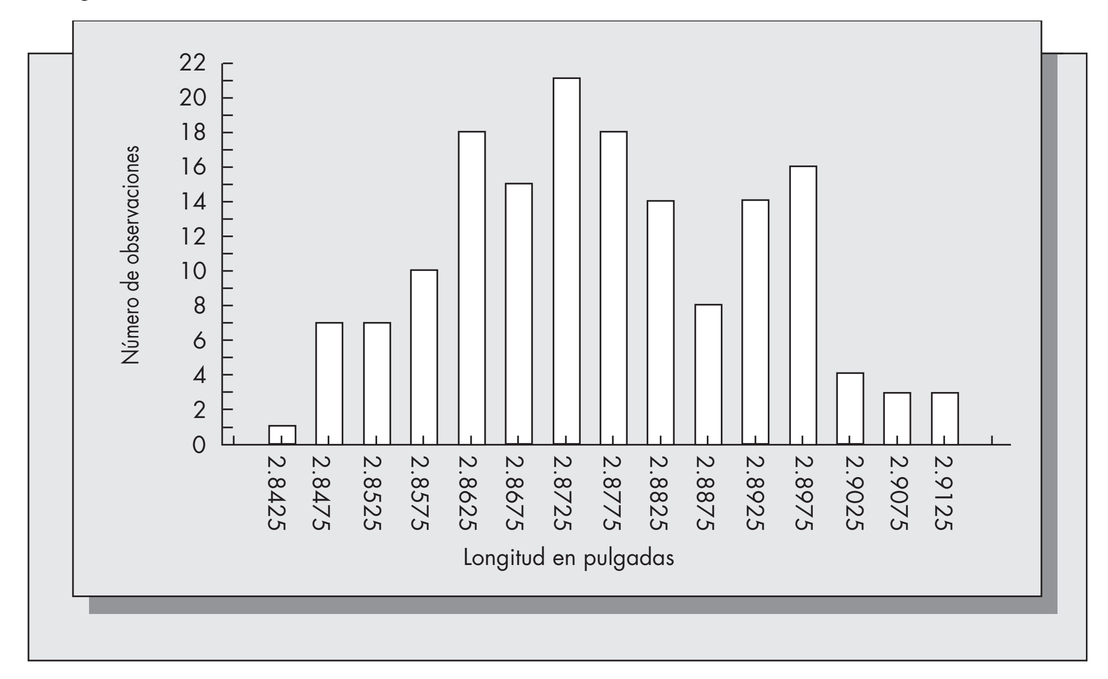
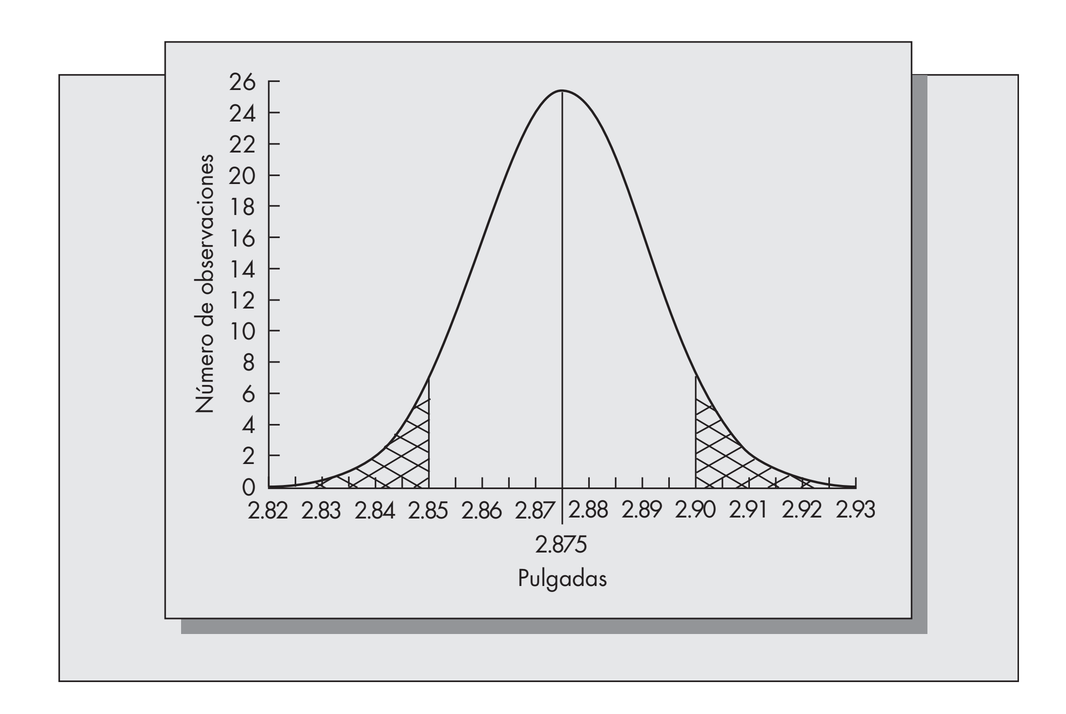


FIGURA 11-6
Curva normal teórica de la longitud del brazo.



Interprete a $X_1, X_2, \ldots, X_n$ como los valores muestrales (de la muestra aleatoria). La media muestral basada en todas las 150 observaciones es 2.875, y la varianza muestral es 0.0002434. La estimación de la desviación estándar de la población es la raíz cuadrada de la varianza muestral, que es 0.0156.

Ahora Howard puede estimar la fracción de los brazos producidos que se sitúan fuera de las tolerancias publicitadas. Un brazo excederá la tolerancia si es mayor que 2.90 pulgadas o menor que 2.85 pulgadas. El área de la región sombreada de la figura 11-6 representa la probabilidad de que esto ocurra, que evidentemente no es despreciable. De hecho,

$$P\{X > 2.90 \text{ or } X < 2.85\} = P\{|X - \mu| > 0.025\}$$
$$= P\left\{\left|\frac{X - \mu}{\sigma}\right| > 1.602\right\}$$
$$= P\{|Z| > 1.602\} = .11.$$

Encontramos la probabilidad, .11, en una tabla de la distribución normal (la tabla A-1 al final de este libro). Así, aproximadamente 11 por ciento de los brazos producidos en los últimos 30 días no cumplen las especificaciones publicadas por la compañía. Sin embargo, las tasas de fallas de los discos debidas a la incompatibilidad de los brazos han sido muy bajas (menores que el 1 por ciento). Howard presentó estos resultados al director de manufactura, quien discutió el problema con el presidente de la compañía y con el director de ingeniería. Tras cierta investigación adicional, Howard concluyó que las tolerancias originales no eran consistentes con los requerimientos de diseño del disco. Se encontró que una tolerancia de ± 0.05 pulgadas sería suficiente. Las tolerancias más estrictas se basaban en un diseño anterior, y el departamento simplemente olvidó revisar sus cifras. Las pruebas posteriores mostraron que una tolerancia de ± 0.05 pulgadas sería mucho mas realista y consistente con la operación del disco. Howard quedó satisfecho con que las tolerancias revisadas incluyeran más del 99 por ciento de la población.

Este ejemplo hace surgir un punto importante en la aplicación de los principios estadísticos al control de la calidad. Si las gráficas de control van a usarse para comparar las características de los artículos manufacturados con requerimientos de diseño previamente especificados, entonces las tolerancias deseadas y la variación estadística observada en la muestra deben ser consistentes. Si las tolerancias son mucho más estrictas que la variación observada en la muestra, como en el ejemplo 11.1, entonces frecuentemente serán sobrepasadas aun cuando el proceso esté bajo control. También puede presentarse la situación opuesta:

las tolerancias observadas pueden ser mucho más *holgadas* que la variación observada en la población. En este caso una observación puede estar fuera de control en relación con otros valores muestrales, pero puede situarse dentro de las tolerancias deseadas. Que el proceso esté bajo control arrojaría poca información acerca de si las partes están cumpliendo con las especificaciones.

Gráficas $\overline{X}$

Considere el ejemplo 11.1. Digamos que Howard decide construir una gráfica $\overline{X}$ para los datos resumidos en la tabla 11-1. Una gráfica $\overline{X}$ requiere que los datos se dividan en subgrupos de tamaño fijo. El tamaño de los subgrupos para el ejemplo es $n = 5$. El tamaño del subgrupo debe ser cuando menos 4 para que sea aplicable el teorema del límite central.

Para construir una gráfica $\overline{X}$, se necesita estimar la media muestral y la varianza muestral de la población. Esto puede hacerse usando las fórmulas dadas. Sin embargo, generalmente no se recomienda que se use la desviación estándar de la muestra como un estimado de σ al construir una gráfica $\overline{X}$. Para que s sea un estimado exacto de σ, es necesario que la media básica de la muestra sea constante. Como el propósito de una gráfica $\overline{X}$ es determinar si ha ocurrido una desviación en la media, no debemos suponer a priori que la media es constante al estimar σ. Un método alternativo para estimar la variación de la muestra que permanece exacta cuando cambia la media poblacional emplea los rangos de los datos. Aun si hay una desviación en la media del proceso, los rangos serán estables siempre que la variación del proceso sea estable. Existe una relación entre la desviación estándar de la población y el rango de los subgrupos de un tamaño dado que depende del tamaño del subgrupo. Es decir, existe una constante d_2 tal que

$$\hat{\sigma} = \frac{\overline{R}}{d_2},$$

donde R es el promedio de los rangos observados y $\hat{\sigma}$ es un estimado de la desviación estándar de la población. Las constantes d_2 para los diferentes tamaños de subgrupo se muestran en la tabla A-5 al final de este libro. Para los datos presentados en la tabla 11-1, el promedio de los 30 rangos resulta ser 0.035756. El valor de d_2 para subgrupos de tamaño 5 es 2.326. Entonces el estimado de σ con base en estos datos es

$$\hat{\sigma} = 0.035756/2.326 = 0.01537,$$

que es bastante cercano al estimado de la desviación estándar usando la desviación estándar de la muestra s.

Dados los estimados de la media y la desviación estándar del promedio del grupo, las gráficas de control se construyen de la siguiente manera: se dibujan líneas para los límites de control superior e inferior en $\overline{X} \pm 3\sigma/\sqrt{n}$. Los promedios de los grupos se grafican con una intermitencia diaria. Se dice que el proceso está fuera de control si una observación se sitúa fuera de los límites de control. La gráfica $\overline{X}$ para la muestra de 30 días para el brazo de seguimiento se muestra en la figura 11-7. Observe que el proceso parece estar bajo control, ya que todas las observaciones están situadas dentro de los límites 3σ.

Relación con la estadística clásica

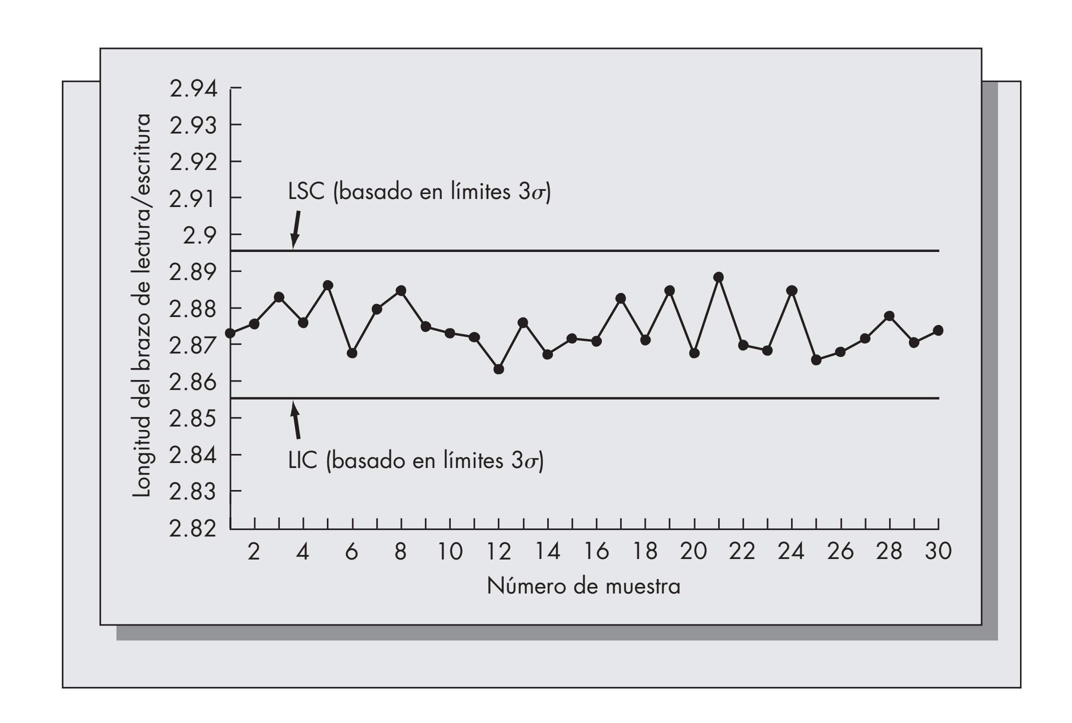
Aquí consideramos las gráficas de control estadístico en el contexto de las pruebas de hipótesis de la estadística clásica. La hipótesis nula es que el proceso subyacente está bajo control. Es decir, tenemos las hipótesis

H_0: el proceso está bajo control.

H_1: el proceso está fuera de control.

Interpretamos la palabra *control* con el significado de que el mecanismo subyacente al azar que genera las observaciones es estable con el tiempo. Para las gráficas $\overline{X}$, probamos si la

FIGURA 11-7
Gráfica $\overline{X}$ para los datos de los brazos de lectura/escritura.



media del proceso ha experimentado alguna desviación. Hay dos maneras de llegar a la conclusión equivocada: rechazar la hipótesis nula cuando es verdadera (concluir que el proceso está fuera de control cuando está bajo control) y rechazar la hipótesis alternativa cuando es verdadera (concluir que el proceso está bajo control cuando está fuera de control). A estos se les llama, respectivamente, los errores Tipo 1 y Tipo 2. Usamos el símbolo α para representar la probabilidad de un error tipo 1 y , para representar la probabilidad de un error tipo 2. Una prueba es una regla que indica cuando rechazar a H_0 con base en los valores muestrales. Una prueba requiere la especificación de un valor aceptable de α. Conceptualmente, estamos haciendo lo mismo cuando usamos las gráficas de control.

La hipótesis de que el proceso esté bajo control se rechaza si un valor observado de $\overline{X}$ se sitúa fuera de los límites de control. Podemos establecer los valores del límite superior de control (LSC) y del límite inferior de control (LIC) con base en la especificación de cualquier valor de α.

$$\begin{aligned}
\alpha &= P\{\text{error Tipo 1}\} \\
&= P\{\text{se observa una señal fuera de control} \mid \text{es proceso está bajo control}\} \\
&= P\{\overline{X} < \text{LIC or } \overline{X} > \text{LSC} \mid \text{la media verdadera es } \mu\} \\
&= P\{\overline{X} < \text{LIC} \mid \mu\} + P\{\overline{X} > \text{LSC} \mid \mu\} \\
&= P\left\{\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < \frac{\text{LIC} - \mu}{\sigma/\sqrt{n}}\right\} + P\left\{\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} > \frac{\text{LSC} - \mu}{\sigma/\sqrt{n}}\right\} \\
&= P\left\{Z < \frac{\text{LIC} - \mu}{\sigma/\sqrt{n}}\right\} + P\left\{Z > \frac{\text{LSC} - \mu}{\sigma/\sqrt{n}}\right\}.
\end{aligned}$$

Como la distribución normal es simétrica, establecemos

$$\frac{\text{LIC} - \mu}{\sigma/\sqrt{n}} = -z_{\alpha/2}, \qquad \frac{\text{LSC} - \mu}{\sigma/\sqrt{n}} = z_{\alpha/2},$$

con lo que se obtiene

$$\text{LSC} = \mu + \frac{\sigma z_{\alpha/2}}{\sqrt{n}},$$

$$\text{LIC} = \mu - \frac{\sigma z_{\alpha/2}}{\sqrt{n}}.$$

Al hacer $z_{\alpha/2} = 3$, obtenemos los populares límites de control tres sigma. Esto es equivalente a seleccionar un valor de $\alpha = .0026$. Este valor específico de α es el que se usa tradicionalmente; no es necesariamente el único que tiene sentido. En algunas aplicaciones, podría desearse incrementar la posibilidad de reconocer cuando el proceso sale de control. Entonces se usaría un valor mayor de ·, que condujera a límites de control más estrictos. Por ejemplo, un valor de α de .05 conduciría a límites dos sigma en vez de límites tres sigma.

Gráficas *R*

La gráfica $\overline{X}$ se usa para probar si hay alguna desviación en el valor medio de un proceso. En muchos casos también estamos interesados en hacer pruebas para ver si hay una desviación en la varianza del proceso. La variación del proceso puede monitorearse examinando las varianzas de la muestra de las observaciones del subgrupo. Sin embargo, los rangos de los subgrupos dan aproximadamente la misma información y son mucho más fáciles de calcular. La teoría detrás de la gráfica R es que cuando la población que se maneja es normal, existe una relación entre el rango de la muestra y la desviación estándar de la muestra que depende del tamaño de la muestra. Si $\overline{R}$ es el promedio de los rangos de todos los subgrupos de tamaño n, entonces, de acuerdo con lo ya estudiado en esta sección, tenemos

$$\hat{\sigma} = \overline{R}/d_2,$$

donde d_2, que depende de n, se muestra en la tabla A-5 al final de este libro.

Normalmente, se desarrollaría una gráfica R antes que una gráfica $\overline{X}$ con objeto de obtener un estimado confiable de la varianza. El estimado $\hat{\sigma}$ es menos sensible a los cambios de la media del proceso que el estimado s. El propósito de la gráfica R es determinar si la variación del proceso resulta estable. Los límites superior e inferior de esta gráfica están dados por las fórmulas

$$\text{LIC} = d_3\overline{R},$$
$$\text{LSC} = d_4\overline{R}.$$

Los valores de las constantes d_3 y d_4 se muestran en la tabla A-6 en la parte posterior de este libro. Los valores dados para estas constantes suponen límites tres sigma para el proceso del rango.

Ejemplo *11.1 (continuación)*

Nuevamente considere los datos del brazo de lectura/escritura en la tabla 11-1. Los rangos de las muestras de tamaño $n = 5$ se muestran en la columna final de la tabla. Como se afirmó anteriormente, el promedio de estos 30 rangos es 0.035756. Éste es el valor de $\overline{R}$ y se convierte en el punto central de la gráfica R. Los límites de control superior e inferior para R se calculan usando las fórmulas dadas y la tabla A-6. Para el caso de $n = 5$, tenemos $d_3 = 0$ y $d_4 = 2.11$, conduciendo así a los siguientes límites de control para la gráfica R de Wonderdisk:

$$\text{LIC} = (0)(0.035756) = 0.$$
$$\text{LSC} = (2.11)(0.035756) = 0.07545.$$

Estos son los límites tres sigma.

La gráfica R de Wonderdisk se ilustra en la figura 11-8. Como todos los valores observados de R se sitúan dentro de los límites de control, la variación del proceso está bajo control.

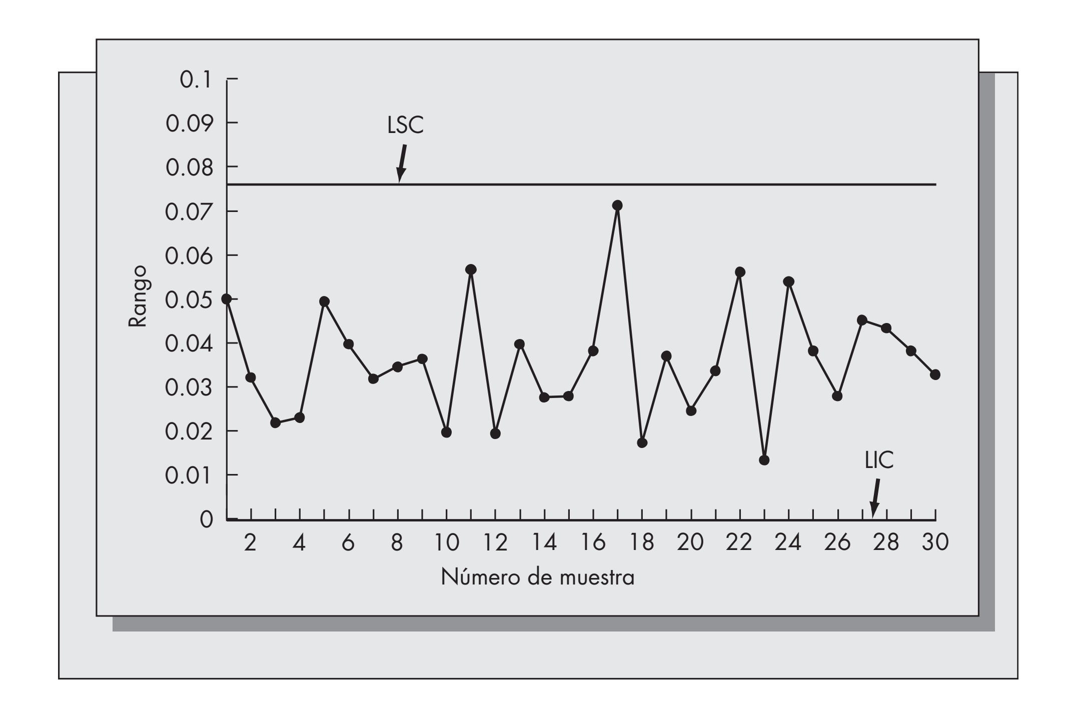


FIGURA 11-8
Gráfica *R* para los datos de los brazos de lectura-escritura.

Las gráficas *R* se usan para probar si hay una desviación en la variación del proceso. Para que sean correctos los valores de los estimados usados al construir la gráfica $\overline{X}$, la varianza del proceso debe ser constante. Es decir, se recomienda que la gráfica *R* se use *antes* que la gráfica $\overline{X}$, ya que una gráfica $\overline{X}$ supone que la variación del proceso es estable.

Las gráficas *R* no son el único medio para probar la estabilidad de la variación del proceso. También podría usarse una gráfica σ. Se grafican las desviaciones estándar de la muestra de los subgrupos con respecto al tiempo para determinar si ocurre una desviación estadísticamente significativa en estos valores y cuándo sucede éste. Las gráficas sigma se usan rara vez en la práctica por dos razones:

1. Cuesta más trabajo calcular las desviaciones estándar de la muestra para cada subgrupo que calcular los rangos.
2. Las gráficas *R* y las gráficas σ casi siempre proporcionan los mismos resultados.

Por estas razones, en este libro no estudiaremos las gráficas σ.

Problemas para la sección 11.2

7. El grupo de control de calidad de una compañía manufacturera está planeando usar gráficas de control para monitorear la producción de cierta parte. Las especificaciones de la parte requieren que cada unidad pese entre 13.0 y 15.5 onzas con un valor meta de 14.25. Una muestra de 75 observaciones conduce a lo siguiente:

$$\sum_{i=1}^{75} X_i = 1\,065, \qquad \sum_{i=1}^{75} X_i^2 = 15\,165.$$

a) ¿Son consistentes las especificaciones con la variación estadística de la muestra? (Sugerencia: use la fórmula de cálculo para la varianza de la muestra:

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right].)$$

b) ¿Qué problemas anticipa usted si el grupo intenta usar las gráficas $\overline{X}$ y R para este proceso?

c) Si las 75 observaciones son estadísticamente estables, ¿qué porcentaje de los artículos manufacturados van a situarse fuera de las tolerancias?

8. Las gráficas de control para $\overline{X}$ y R se conservan para la resistencia al corte de las soldaduras de punto. Cien observaciones divididas en subgrupos de tamaño cinco se usan como línea base para construir las gráficas, y se calculan estimaciones de μ y σ a partir de estas observaciones. Suponga que las 100 observaciones son $X_1, X_2, \ldots, X_{100}$ y los rangos de los 20 subgrupos son $R_1, R_2, \ldots, R_{20}$. A partir de estos datos de la línea base se calculan las siguientes cantidades:

$$\sum_{i=1}^{100} X_i = 97\,500, \qquad \sum_{j=1}^{20} R_j = 1\,042.$$

Usando esta información, calcule los valores de los límites tres sigma para ambas gráficas.

9. Para el problema 8, suponga que la probabilidad de concluir que el proceso está fuera de control cuando realmente está bajo control se establece como .02. Encuentre los límites de control superior e inferior para la gráfica $\overline{X}$ resultante.

10. Suponga que la media del proceso se desvía a 1 011 en el problema 8.

a) ¿Cuál es la probabilidad de que la gráfica $\overline{X}$ *no* indique una condición fuera de control a partir de un muestreo simple de un subgrupo? (éste es precisamente el error Tipo 2.)

b) ¿Cuál es la probabilidad de que la gráfica $\overline{X}$ no indique una condición fuera de control después de muestrear 20 subgrupos?

11. Un servicio de procesamiento de películas monitorea la calidad del proceso de revelado con equipo sensible a la luz. La medición de la exactitud es un número con un valor meta de cero. Suponga que una gráfica X con subgrupos de tamaño cinco se usa para monitorear el proceso y los límites de control son LSC = 1.5 y LIC = −1.5. Suponga que el estimado de la media del proceso es cero y para la desviación estándar del proceso es 1.30.

a) ¿Cual es el valor de · para esta gráfica de control?

b) Encuentre los valores de LSC y LIC con base en los límites tres sigma.

c) Suponga que la media del proceso se desvía a 1. ¿Cuál es la probabilidad de que la desviación se detecte en el primer subgrupo después que ocurra la desviación?

12. Una gráfica R se usa para monitorear la variación de los pesos de los paquetes de galletas con chispas de chocolate producidas por una industria nacional grande de alimentos horneados. Un analista ha recopilado una línea base de 200 observaciones para construir la gráfica. Suponga un valor calculado de $\overline{R}$ de 3.825.

a) Si se usarán subgrupos de tamaño seis, calcule el valor de los límites tres sigma para esta gráfica.

b) Si se usa una gráfica $\overline{X}$ basada en los límites tres sigma, ¿Cuál es la diferencia entre el LSC y el LIC?

13. Un proceso se monitorea con el uso de una gráfica $\overline{X}$ con LSC = 13.8 y LIC = 8.2. La desviación estándar del proceso se estima como 6.6. Si la gráfica X se basa en los límites tres sigma,

a) ¿Cuál es el estimado de la media del proceso?

b) ¿Cuál es el tamaño de cada uno de los subgrupos de muestreo?

11.3 GRÁFICAS DE CONTROL PARA ATRIBUTOS: LA GRÁFICA *p*

Las gráficas $\overline{X}$ y R son herramientas valiosas para controlar los procesos cuando la salida del proceso pueda expresarse como una sola variable real. Esto es apropiado si existe una dimensión de una cualidad individual como la longitud, el ancho, o la dureza. Hay dos circunstancias en las cuales las gráficas de control para variables no son apropiadas: (1) cuando la preocupación es si el artículo tiene un atributo específico (por ejemplo, el problema puede ser si el artículo funciona) y (2) cuando hay muchas variables de calidad diferentes. En el caso (2) no es práctico ni es efectivo en los costos mantener gráficas de control separadas para cada variable. El artículo tiene los atributos deseados o no los tiene.

Si se usan las gráficas de control para atributos, cada valor muestral es 1 o 0. Un 1 significa que el artículo es aceptable, y un 0 significa que no lo es. Sea n el tamaño del subgrupo de muestreo y defina a la variable aleatoria X como el número total de elementos defectuosos en el subgrupo. Supondremos que cada subgrupo representa un muestreo de la producción de un día. La teoría sería exactamente la misma independientemente de que el intervalo de muestreo sea de una hora, un día o un mes. Ya que X cuenta el número de elementos defectuosos en un tamaño fijo de muestra, la distribución subyacente de X es binomial con parámetros n y p. Interprete a p como la proporción de elementos defectuosos producidos y n como el número de artículos muestreados en cada grupo (comúnmente, n es el número de artículos muestreados cada día). Una gráfica p se usaría para determinar si existe una desviación importante en el valor verdadero de p.

Aunque podrían construirse gráficas p con base en la distribución binomial exacta, es más común usar una aproximación normal. Además, como nuestro interés reside en estimar el valor de p, rastreamos la variable aleatoria X/n, cuya esperanza es p, en vez de X misma. Es fácil demostrar que

$$E(X/n) = p,$$
$$\text{Var}(X/n) = p(1 - p)/n.$$

Para valores grandes de n, el teorema del límite central nos dice que X/n tiene una distribución aproximadamente normal con parámetros $\mu = p$ y $\sigma = \sqrt{p(1 - p)/n}$. Con el uso de una aproximación normal, los límites tres sigma tradicionales son

$$\text{LSC} = p + 3\sqrt{\frac{p(1 - p)}{n}},$$

$$\text{LIC} = p - 3\sqrt{\frac{p(1 - p)}{n}}.$$

El estimado de p, la proporción verdadera de artículos defectuosos en la población, es p, la fracción promedio de artículos defectuosos observados durante un periodo razonable de la línea base. Se dice que el proceso está bajo control siempre que la fracción observada de artículos defectuosos para cada subgrupo permanezca dentro de los límites de control superior e inferior.

Ejemplo *11.2* Xezet, un fabricante de discos flexibles, inspecciona una muestra de 50 discos de la producción diaria. Con base en diferentes atributos, el inspector de calidad clasifica a cada disco como aceptable o no. La experiencia para un periodo típico de 20 días se resume en la tabla 11–2.

Para construir una gráfica de control para la fracción de artículos defectuosos, se necesita tener un estimado exacto de la fracción verdadera de artículos defectuosos en la población completa, p. Con base en los datos de la tabla 11-2 construimos una gráfica de control preliminar para determinar si los datos de la línea base están bajo control. El número total de artículos defectuosos observados durante los 20

TABLA 11-2 **Número de discos rechazados**

Fecha	Número de rechazos	Fecha	Número de rechazos
3/18	3	4/1	0
3/19	10	4/2	4
3/20	13	4/3	9
3/21	4	4/4	22
3/22	12	4/5	7
3/25	14	4/8	6
3/26	8	4/9	18
3/27	7	4/10	3
3/28	19	4/11	9
3/29	1	4/12	7

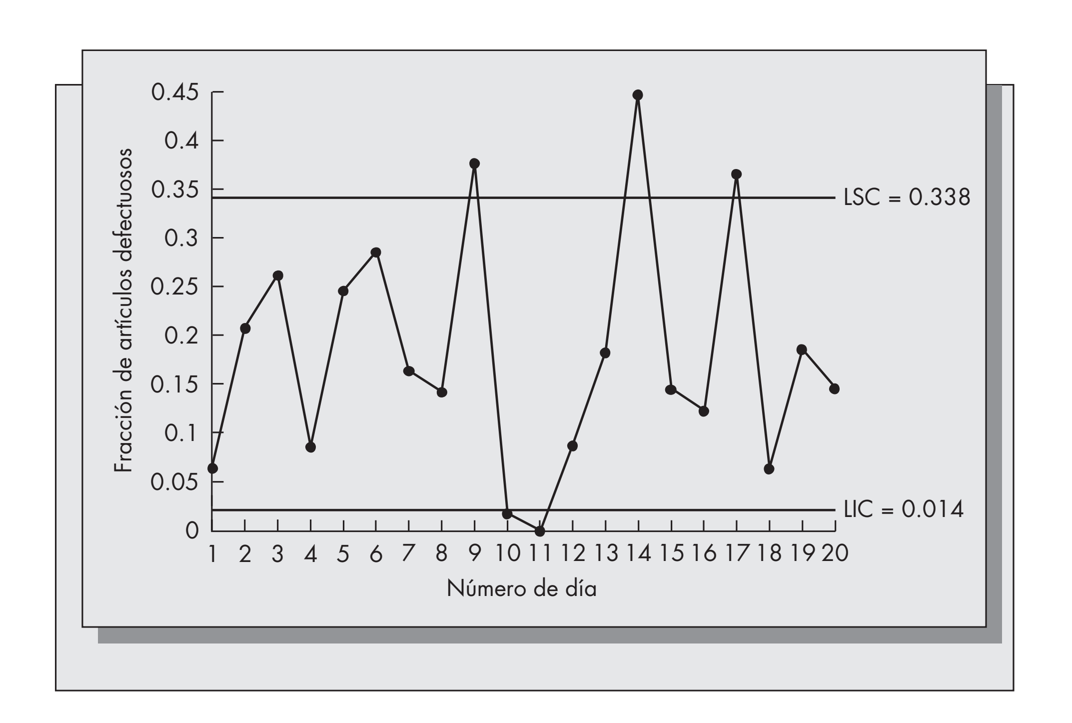


FIGURA 11-9
Gráfica *p* preliminar para los datos de los discos flexibles Xezet (véase al ejemplo 11.2).

días es 176. La producción total para el mismo periodo de tiempo es 1 000 discos. Así, el estimado actual de la proporción de artículos defectuosos es 176/1 000 = .176. El estimado actual de σ es

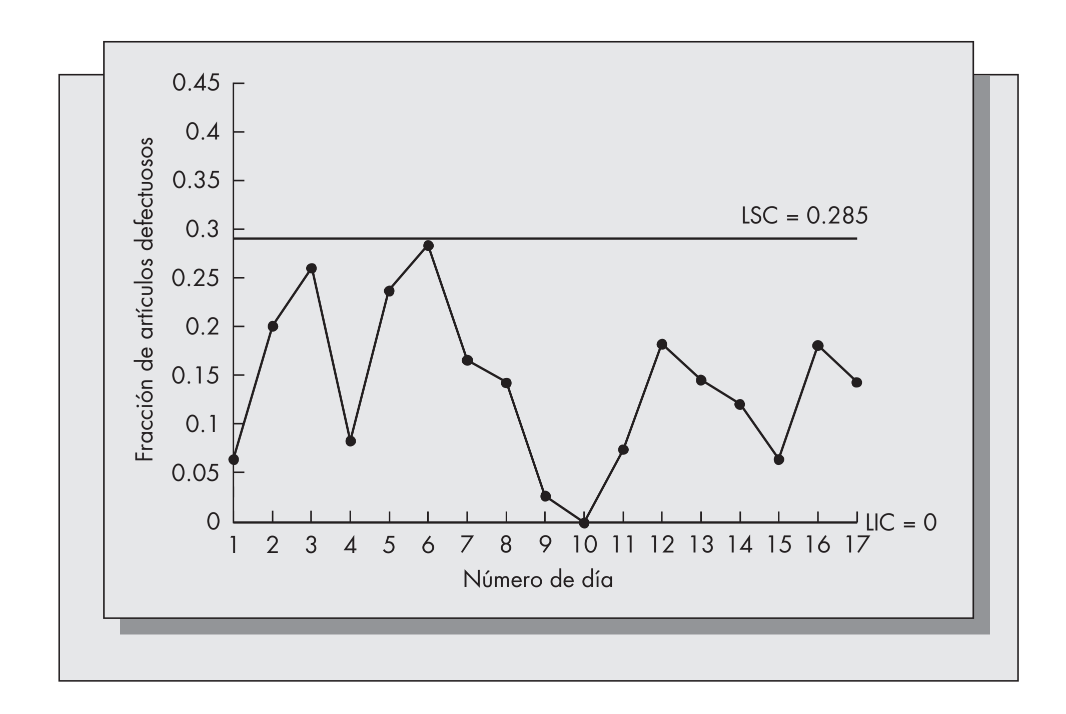
$$\hat{\sigma} = \sqrt{\frac{\overline{p}(1-\overline{p})}{n}} = \sqrt{\frac{(.176)(.824)}{50}} = .054.$$

Con base en los límites tres sigma obtenemos

$$\text{LIC} = 0.176 + (3)(0.054) = 0.338,$$
$$\text{LSC} = 0.176 - (3)(0.054) = 0.014.$$

La figura 11-9 es la gráfica preliminar de control para la fracción de artículos defectuosos. Observe que cuatro puntos están fuera de control. Estos corresponden a los días 9 (3/28), 11 (4/1), 14 (4/4), y 17 (4/9). No debemos preocuparnos por un punto que se sitúa debajo del límite de control inferior, porque eso muestra una tasa de artículos defectuosos mejor que lo esperado. El gerente de producción considera a los tres puntos restantes fuera de control y se percata de que corresponden a tres días en los cuales un empleado clave estuvo ausente del trabajo por razones personales. El trabajo del empleado

FIGURA 11-10
Gráfica p revisada para los datos de los discos flexibles Xezet (véase el ejemplo 11.2).



requiere algún equipo muy complejo de galvanoplastia. La elevada tasa de artículos defectuosos en estos días fue evidentemente el resultado de la falta de experiencia temporal del empleado con el equipo de galvanoplastia.

Como los puntos fuera de control fueron explicados por causas asignables, eliminamos estos puntos de la muestra. Los datos de la línea base consisten ahora en los datos listados en la tabla 11-2 eliminando los tres días correspondientes a los puntos fuera de control. Es decir, la base de datos ahora consiste de un total de 17 días. Ahora recalculamos el estimado de p con base en los datos revisados. Obtenemos

$$\overline{p} = 117/(17)(50) = .138$$

y

$$\hat{\sigma} = \sqrt{(.138)(.862)/50} = 0.049.$$

Con base en estos nuevos estimados para p y σ obtenemos

$$\text{LIC} = 0.138 + (3)(0.049) = 0.285,$$
$$\text{LSC} = 0.138 - (3)(0.049) = -0.009.$$

Si LIC < 0, hacemos LIC $= 0$, porque es imposible observar un valor de p que sea negativo. En la figura 11-10 trazamos la gráfica *p* revisada para los datos de los discos flexibles. Observe que con los tres puntos fuera de control eliminados, ahora todos los puntos restantes se sitúan dentro de los límites de control. No siempre se dará este caso, puesto que los límites de control inferior y superior son más cercanos en la figura 11-10 que lo que fueron en la figura 11-9. Ahora se usarían los límites de control corregidos para monitorear las futuras observaciones de la fracción de artículos defectuosos para este proceso.

Gráficas *p* para tamaños variables de subgrupo

En la sección 11.3 supusimos que el número de artículos inspeccionados en cada subgrupo era el mismo. Este supuesto es razonable cuando los subgrupos se muestrean periódicamente de lotes grandes. Sin embargo, en muchas circunstancias se producen pocos artículos cada día y en

TABLA 11-3 Artículos defectuosos observados en tornos industriales (véase el ejemplo 11.3)

Número de día	Producción	Número de artículos defectuosos	Valor *Z* estandarizado
1	82	8	−0.4451
2	76	12	1.2320
3	85	6	−1.2382
4	53	5	−0.4319
5	30	3	−0.2270
6	121	14	0.0893
7	63	11	1.5404
8	80	9	−0.0178
9	88	7	−0.9946
10	97	8	−0.9532
11	91	13	0.8953
12	77	14	1.9029
13	71	6	−0.7614
14	95	8	−0.8899
15	102	13	0.4566

consecuencia se sujetan a una inspección al 100 por ciento. Si la producción diaria varía, el tamaño del subgrupo también variará. Aquí basaremos el análisis en la variable estandarizada Z:

$$Z = \frac{p - \bar{p}}{\sqrt{\bar{p}(1 - \bar{p})/n}},$$

que es aproximadamente normal estándar e independiente de n. Los límites de control inferior y superior se establecerían en -3 y $+3$, respectivamente, para obtener límites tres sigma, y la gráfica de control monitorearía valores sucesivos de la variable estandarizada Z.

Ejemplo *11.3*

Un productor de tornos industriales ejecuta la inspección y la prueba al 100 por ciento de la producción diaria. La producción de tornos varía diariamente con base en las órdenes que se esperan y en el calendario de producción para las otras máquinas que la compañía produce. El número observado de tornos defectuosos y la producción diaria durante los últimos 15 días se dan en la tabla 11-3.

Se calculan los valores de Z normalizados en la última columna de la siguiente manera. Primero, estime p a partir de la historia completa de 15 días calculando el cociente del número total de artículos defectuosos observados entre el número total de artículos producidos. Para los datos de la tabla 11-3 obtenemos $\bar{p} = 137/1{,}211 = .1131$. Los valores de Z se calculan ahora con la fórmula dada. Por ejemplo, para el día 1

$$Z = \frac{8/82 - 137/1{,}211}{\sqrt{\dfrac{(137/1{,}211)(1 - 137/1{,}211)}{82}}} = -0.4451.$$

Como los límites tres sigma para Z son simplemente ± 3, es evidente de la tabla 11-3 que este proceso está bajo control.

Problemas para la sección 11.3

14. Un fabricante de película fotográfica produce en promedio 12,000 rollos de diferentes tipos de película cada día. Un inspector muestrea 25 de la producción diaria y los prueba para valorar la exactitud de la reproducción a colores y la calidad general de la película.

Durante 23 días de producción consecutiva, el inspector rechazó las siguientes cantidades de rollo de película:

Día	Cantidad rechazada	Día	Cantidad rechazada
1	2	13	1
2	3	14	1
3	2	15	0
4	1	16	0
5	3	17	0
6	4	18	2
7	7	19	3
8	1	20	1
9	3	21	2
10	3	22	2
11	0	23	1
12	0		

a) Con base en estas observaciones, calcule los límites tres sigma para una gráfica *p*.

b) ¿Se sitúa cualquiera de estos puntos fuera de los límites de control? Si así es, recalcule los límites tres sigma después de eliminar los puntos fuera de control.

15. Applied Machines produce equipo de prueba grande para circuitos integrados. Las máquinas se fabrican bajo pedido, de modo que la tasa de producción varía de mes en mes. Antes de ser embarcada, cada máquina se sujeta a pruebas extensas. Con base en las pruebas, la máquina pasa o se regresa para reprocesarla. Durante los últimos 20 meses la compañía ha tenido que reprocesar las siguientes cantidades de máquinas:

Mes	Número producido	Número reprocesados	Mes	Número producido	Número reprocesados
1	23	3	11	17	3
2	28	3	12	4	0
3	16	1	13	14	2
4	6	0	14	0	0
5	41	2	15	18	6
6	32	4	16	0	0
7	29	5	17	33	4
8	19	2	18	46	5
9	12	1	19	21	7
10	7	1	20	29	7

Determine si el proceso estuvo en control para el periodo de 20 meses usando una versión estandarizada de la gráfica *p*. Suponga límites tres sigma para la gráfica de control. (Sugerencia: ¿realmente tiene 20 meses de datos?)

16. Considere el ejemplo de Applied Machines presentado en el problema 15. Con base en la estimación de la probabilidad de que una máquina se regrese para su reproceso, calculada a partir de los 20 meses de datos, determine lo siguiente:

a) Si la compañía produce 35 máquinas en un mes específico, ¿cuántas, en promedio, necesitan reprocesarse?

b) De 100 máquinas producidas, ¿cuál es la probabilidad de que más del 20 por ciento de éstas necesiten reprocesarse? (Use la aproximación normal a la binomial para sus cálculos. Se estudia en el apéndice A-11.)

17. Para un periodo de 12 horas consecutivas de producción, las muestras de tamaño 50 condujeron a las siguientes proporciones de artículos defectuosos:

Número de muestra	Proporción de artículos defectuosos	Número de muestra	Proporción de artículos defectuosos
1	.04	7	.10
2	.02	8	.10
3	.06	9	.06
4	.08	10	.08
5	.08	11	.04
6	.04	12	.04

a) ¿Cuáles son los límites de control tres sigma para este proceso?

b) ¿Se sitúa cualquiera de los puntos muestrales fuera de los límites de control?

c) La compañía garantiza una tasa de defectos de 3 por ciento en estos artículos. ¿Son consistentes las proporciones observadas con un valor meta de 3 por ciento de artículos defectuosos? ¿Qué dificultad surgiría si los límites de control se basaran en un valor meta de 0.03? En vista de la garantía de la compañía, ¿qué dificultad surgiría si se usaran los límites de control calculados en el inciso *a*?

11.4 LA GRÁFICA *c*

Hay otras gráficas de control además de las gráficas $\overline{X}$, R y p. Aunque la forma de la distribución de la variable aleatoria apropiada depende de la aplicación, el enfoque básico es el mismo. En general, debe determinarse la distribución de probabilidades de la variable aleatoria de interés, y encontrar los límites de control superior e inferior que contienen al universo de observaciones con un nivel deseado de confianza. Generalmente se establece que la probabilidad de situarse fuera de los límites de control debe ser menor que .01.

La gráfica p es apropiada para clasificar un artículo como bueno o malo. Sin embargo, frecuentemente nos preocupa cuántos defectos tiene un artículo o un grupo de artículos. Un artículo es aceptable si el número de sus defectos no es demasiado grande. Por ejemplo, un refrigerador que tenga unos cuantos rasguños podría considerarse aceptable, pero uno que tenga demasiados rasguños podría considerarse inaceptable. Como otro ejemplo, para una fábrica textil que manufactura telas, tanto el fabricante como el consumidor se interesarían en el número de defectos por yarda de tela.

La *gráfica c* se basa en la observación de que si los defectos se presentan en forma completamente aleatoria, entonces la distribución de probabilidad del número de defectos por unidad de producción tiene la distribución de Poisson. Si c representa el número medio verdadero de los defectos en una unidad de producción, entonces la posibilidad de que haya k defectos en una unidad es

$$P\{\text{Número de defectos en una unidad} = k\} = \frac{e^{-c}c^k}{k!} \qquad \text{para } k = 0, 1, 2, \ldots.$$

Al usar una gráfica de control para el número de defectos, el tamaño de la muestra debe ser el mismo para cada inspección. Se estima el valor de c a partir de los datos de la línea base calculando la media muestral del número observado de defectos por unidad de producción. Si $c \geq 20$, la distribución normal proporciona una aproximación razonable a la de Poisson. Como

la media y la varianza de la distribución Poisson son ambas iguales a c, se concluye que para c grande,

$$Z = \frac{X - c}{\sqrt{c}}$$

es aproximadamente normal estándar. Usando los límites tres sigma tradicionales, los límites de control superior e inferior para la gráfica c son

$$\text{LIC} = c - 3\sqrt{c},$$
$$\text{LSC} = c + 3\sqrt{c}.$$

Se desarrolla y se usa la gráfica c de la misma manera que las gráficas $\overline{X}$, R, y p.

Ejemplo *11.4*

Leatherworks produce diferentes artículos de cuero en su planta de Montpelier, Vermont. Las inspecciones de las últimas 20 unidades de un portafolio de cuero revelaron los siguientes números de defectos:

Número de unidad	Número de defectos observados	Número de unidad	Número de defectos observados
1	4	11	2
2	3	12	3
3	3	13	6
4	0	14	1
5	2	15	5
6	5	16	4
7	4	17	1
8	2	18	1
9	3	19	2
10	3	20	2

La mayoría de los defectos son el resultado de marcas naturales en el cuero, pero aun así, la compañía no quiere embarcar productos que tengan demasiados defectos en el cuero. Usando estos 20 puntos de datos como una línea base, determine los límites de control superior e inferior que incluyan al universo de las observaciones con probabilidad de .95. ¿Qué límites de control resultan al usar una aproximación normal a la de Poisson?

Solución

Para estimar c calculamos la media muestral de los datos, que se encuentra sumando el número total de los defectos observados y dividiendo entre el número de observaciones. Esto es, $c = 56/20 = 2.8$. Para asegurarse de que una gráfica c es apropiada aquí, debemos hacer una prueba de bondad de ajuste de estos datos para una distribución de Poisson con parámetro 2.8. Se deja al lector la verificación de que los datos realmente se ajusten a una distribución de Poisson. (Las pruebas de bondad de ajuste se describen en casi todos los libros de estadística. La más común es la prueba ji cuadrada.)

Para determinar los límites de control exactos, usamos la tabla A-3 en la parte posterior del libro. Ya que la de Poisson es una distribución discreta, es muy poco probable que podamos encontrar límites de control que contengan exactamente el 95 por ciento de la probabilidad. De la tabla vemos que la probabilidad de que el número de defectos sea menor que o igual que cero es $1.0000 - .9392 = .0608$, que es demasiado grande. Entonces, estableceremos el límite de control inferior igual a cero. Por simetría, el límite superior debe corresponder a una cola derecha de aproximadamente .025, que ocurre cuando $k = 7$. Así, recomendaríamos límites de control de LIC $= 0$ y LSC $= 7$.

Para una distribución normal, aproximadamente dos desviaciones estándar desde la media caen dentro del 95 por ciento de la probabilidad. Así, los límites de control basados en una aproximación normal de la distribución Poisson son

$$\text{LIC} = c - 2\sqrt{c} = 2.8 - (2)\sqrt{2.8} = -0.55 \qquad \text{(se iguala a cero)},$$
$$\text{LSC} = c + 2\sqrt{c} = 2.8 + 2\sqrt{2.8} = 6.2.$$

La aproximación normal no es muy exacta en este caso porque c es demasiado pequeña.

Problemas para la sección 11.4

18. Amertron produce cableado eléctrico en rollos de 100 pies. El proceso de inspección de calidad incluye la selección aleatoria de rollos de alambre y el conteo del número de defectos en cada rollo. Los últimos 20 rollos examinados revelaron los siguientes números de defectos:

Rollo	Número de defectos	Rollo	Número de defectos
1	4	11	2
2	6	12	5
3	2	13	5
4	4	14	7
5	1	15	4
6	9	16	8
7	5	17	6
8	5	18	4
9	3	19	6
10	3	20	4

a) Si el número de defectos por rollo de alambre de 100 pies sigue una distribución de Poisson, ¿cuál es el estimado de c obtenido de estas observaciones?

b) Usando una aproximación normal a la de Poisson, ¿cuáles son los límites de control tres sigma que usted usaría para monitorear este proceso?

c) ¿Están las 20 observaciones dentro de los límites de control?

19. Amertron, estudiado en el problema 18, ha establecido una política de aprobar rollos de alambre que tengan cinco o menos defectos.

a) Con base en la distribución exacta de Poisson, ¿cuál es la proporción de los rollos que aprueban la inspección? (Véase la tabla A-3 al final de este libro.)

b) Estime la respuesta al inciso *a* usando la aproximación normal a la distribución Poisson.

20. Un gran productor nacional de galletas y alimentos horneados usa una gráfica c para monitorear el número de chispas de chocolate en sus galletas de chispas de chocolate. A la compañía le gustaría tener un promedio de seis chispas por galleta. Se muestrea una galleta cada hora. Los resultados de las últimas 12 horas fueron

Hora	Número de chispas por galleta	Hora	Número de chispas por galleta
1	7	7	3
2	4	8	6
3	3	9	3
4	3	10	2
5	5	11	4
6	4	12	4

a) Suponiendo un valor meta de $c = 6$, ¿cuáles son los límites de control superior e inferior para una gráfica c?

b) ¿Son consistentes las 12 observaciones con un valor meta de $c = 6$? Si esas 12 observaciones constituyen una línea base, ¿cuáles son los límites de control superior e inferior que resultan? (Use la aproximación normal para sus cálculos.)

21. Para la compañía mencionada en el problema 20, un comprador de una bolsa de galletas de chispas de chocolate descubre una galleta que no tiene chispas y acusa a la compañía de publicidad fraudulenta. Suponga que la compañía produce 300 000 galletas por año. Si el número esperado de chispas por galleta es de seis, ¿Cuántas galletas horneadas al año no tendrían chispas? Véase la tabla A-3 al final del libro.

11.5 MÉTODOS ESTADÍSTICOS CLÁSICOS Y GRÁFICAS DE CONTROL

Las gráficas de control señalan observaciones no representativas en una muestra. La hipótesis de que ha ocurrido una desviación del proceso también puede probarse mediante los métodos estadísticos clásicos. Por ejemplo, considere la gráfica p. Al construir la gráfica p estamos probando la hipótesis de que ha ocurrido una desviación en el valor subyacente de p, la proporción verdadera de artículos defectuosos en el lote. También puede usarse una tabla de contingencias de $2 - n$ para probar si p ha cambiado. Una variable es el tiempo, y la otra variable es la proporción de artículos defectuosos observados. Se usaría la prueba χ^2 para probar si existe o no una relación entre las dos variables; es decir, si la proporción de artículos defectuosos cambia con el tiempo.

No es necesariamente cierto que la prueba χ^2 dé los mismos resultados que una gráfica p. En general, la prueba χ^2 recomendará que se rechace la hipótesis de que los datos son homogéneos con base en el promedio de las desviaciones de la media estimada, y la gráfica de control recomendará que se rechace la hipótesis de que el proceso está bajo control con base en una desviación grande de una observación individual. Es importante entender esta diferencia para determinar cuál sería el procedimiento más apropiado. Probablemente sea cierto que en el contexto de la manufactura hay mayor preocupación con las desviaciones extremas de unas cuantas observaciones que con el promedio de muchas desviaciones, lo que proporciona una razón para que los profesionales prefieran la metodología de la gráfica de control. Otra razón de la preferencia por las gráficas de control es que son fáciles de usar y de entender. Los gerentes de control de calidad están más familiarizados con las gráficas de control que con los métodos estadísticos clásicos.

Problema para la sección 11.5

22. Considere los datos presentados en el problema 14. El problema 14 requería probar si el proceso se mantenía bajo control mediante el uso de una gráfica p. Pruebe la hipótesis de que el valor de p sea el mismo cada día usando métodos estadísticos clásicos. Es decir, pruebe la hipótesis

$$\mathrm{H}_0\colon p_1 = p_2 = \cdots = p_k$$

contra

$$\mathrm{H}_1\colon \text{No todas las } p_i \text{ son iguales,}$$

donde k es el número de días en el conjunto de datos ($k = 23$ en este caso) y p_i es la proporción verdadera de artículos defectuosos para el día i. Defina a x_i como el número de rollos de película rechazados en el día i y p' como el estimado de p obtenido a partir de los datos [p' se calculó en el problema 14(a)]. El estadístico de prueba está dado por la fórmula

$$\chi^2 = \sum_{i=1}^{k} \frac{(x_i - np')^2}{np'(1 - p')},$$

donde n es el número de artículos muestreados cada día ($n = 25$ en este caso). La prueba es rechazar H_0 si $\chi^2 > \chi^2_{\alpha,k-1}$, donde $\chi^2_{\alpha,k-1}$ es un número obtenido de una tabla de la distribución χ^2. Para $\alpha = .01$ (que es mayor que el valor de α que arroja límites 3σ en una

gráfica p), $\chi^2_{.01,22} = 40.289$. Con base en este valor de α, ¿indica la prueba χ^2 que este proceso está fuera de control? Si la respuesta que usted obtuvo es diferente de la del problema 14(b), ¿cómo explica usted la discrepancia? ¿Qué método es probablemente más adecuado para esta aplicación?

*11.6 DISEÑO ECONÓMICO DE GRÁFICAS $\overline{X}$

El diseño de una gráfica $\overline{X}$ requiere que se determinen diferentes parámetros. Estos incluyen la cantidad de tiempo que transcurre entre el muestreo, el tamaño de la muestra extraída en cada intervalo, y los límites de control superior e inferior. Las penalidades asociadas con los límites de control superior e inferior se reflejan en los errores Tipo 1 y Tipo 2. En esta sección se incorporarán los costos explícitos de estos errores en el análisis, así como los costos de muestreo, y se considera el problema del diseño de una gráfica $\overline{X}$ con base en la minimización de los costos.

El modelo tratado aquí no incluye al intervalo de muestreo como una variable de decisión. En muchas circunstancias el intervalo de muestreo se determina tomando en cuenta consideraciones diferentes del costo. Existen intervalos de tiempo convenientes o naturales basados en la naturaleza del proceso, los artículos que se producen, o las restricciones de personal.

Consideraremos los siguientes tres costos:

1. El costo de muestreo.
2. El costo de búsqueda.
3. El costo de operar fuera de control.

1. *El costo de muestreo*. Suponemos que se muestrean exactamente n artículos en cada periodo. En la mayoría de los casos, el muestreo requiere el tiempo del trabajador, por lo que se incurre en costos de personal. También puede haber costos asociados con el equipo requerido para el muestreo. Además, el muestreo puede requerir pruebas destructivas, añadiendo el costo del artículo mismo. Supondremos que para cada artículo muestreado, hay un costo de a_1. Se concluye que el costo de muestreo incurrido en cada periodo es $a_1 n$.

2. *El costo de búsqueda*. Cuando se señala una condición fuera de control, se supone que existe una causa que la origina. La búsqueda de la causa asignable generalmente requerirá que el proceso se detenga. Si ocurre una señal de fuera de control, hay dos posibilidades: ya sea que el proceso realmente esté fuera de control o que sea una falsa alarma. En cualquiera de los dos casos, supondremos que hay un costo a_2 en el cual se incurre cada vez que se requiere asignar una causa a la condición fuera de control. El costo de búsqueda podría incluir los costos de cerrar la instalación, el tiempo hábil requerido para identificar la causa de la señal, el tiempo requerido para determinar si la señal de fuera de control fue una falsa alarma, y el costo de probar y posiblemente ajustar el equipo. Observe que el costo de búsqueda es probablemente una variable aleatoria: tal vez no sea posible predecir el grado de esfuerzo requerido para buscar una causa imputable a la señal de fuera de control. Si ése es el caso, interprete a_2 como el costo de búsqueda *esperado*.

3. *Operación fuera de control*. El tercer y último costo que consideraremos es el costo de operar el proceso después que se ha salido de control. Existe una mayor posibilidad de producir artículos defectuosos si el proceso está fuera de control. Si se descubren artículos defectuosos durante la inspección, éstos serían convertidos en chatarra o reparados en un tiempo futuro. Una consecuencia aún más seria es que un artículo defectuoso se haga parte de un subensamble más grande, el cual deberá desensamblarse o convertirse en chatarra. Finalmente, los artículos defectuosos pueden encontrar su camino hasta el mercado, lo que conduce a posibles costos de las demandas por garantía, juicios de responsabilidad, y la insatisfacción general de los clientes. Suponga que hay un costo a_3 cada vez que el proceso se opera en una condición fuera de control.

Consideramos solamente el diseño económico de las gráficas $\overline{X}$. Suponga que la media del proceso es μ y la desviación estándar del proceso es σ. Se supone que existe una historia

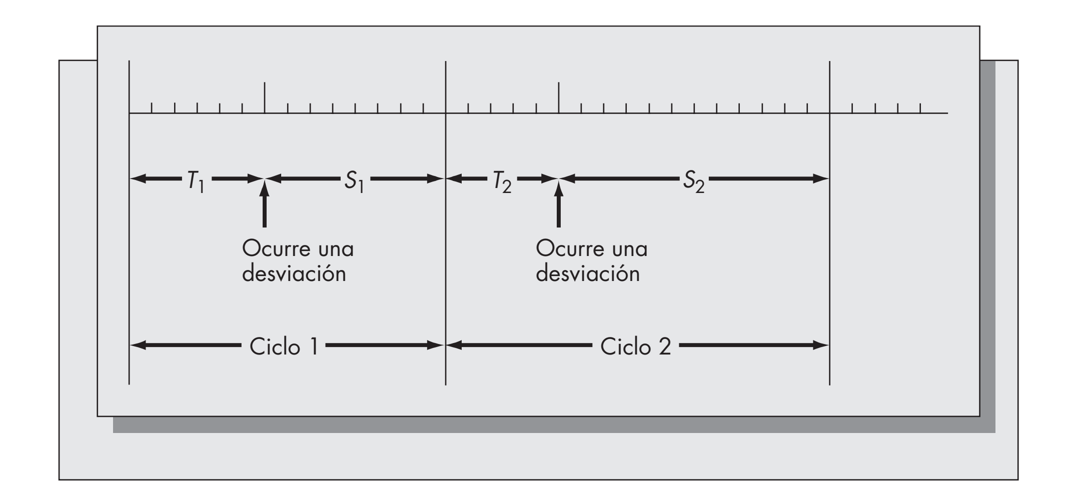


FIGURA 11-11
Ciclos sucesivos en el monitoreo del proceso.

suficiente de observaciones como para que μ y σ puedan estimarse con exactitud. También suponemos que una condición *fuera de control* implica que la media subyacente experimenta una desviación de μ a $\mu + \delta\sigma$ o a $\mu - \delta\sigma$. Así, *fuera de control* significa que la media se desvía en ‰ desviaciones estándar.

Defina un ciclo como el intervalo de tiempo desde el inicio de la producción inmediatamente después de un ajuste, hasta la detección y la eliminación de la causa asignable de la siguiente condición fuera de control. Un ciclo consiste de dos partes. Defina T como el número de periodos que el proceso permanece bajo control siguiendo directamente a un ajuste y S como el número de periodos que el proceso permanece fuera de control hasta que se hace una detección. Un ciclo es la suma $T + S$. Los ciclos sucesivos se ilustran en la figura 11-11. Observe que tanto T como S son variables aleatorias, de modo que la longitud de cada ciclo también es una variable aleatoria. La distribución de probabilidad de T se da subsecuentemente.

Se supone que la gráfica $\overline{X}$ se construye usando los siguientes límites de control:

$$\text{LSC} = \mu + \frac{k\sigma}{}.$$

$$\text{LIC} = \mu - \frac{k\sigma}{}.$$

A lo largo de este capítulo hemos supuesto que $k = 3$, pero esto no siempre puede ser óptimo. El objetivo del análisis de esta sección es determinar los valores económicamente óptimos tanto de k como de n. El método de análisis es determinar una ecuación para el costo global incurrido en un ciclo y otra para la longitud esperada de cada ciclo. En el espíritu del proceso regenerativo (véase Ross, 1970, por ejemplo), tenemos el resultado de que

$$E(\text{Costo por unidad de tiempo}) = \frac{E(\text{Costo por ciclo})}{E(\text{Longitud del ciclo})}.$$

Después de determinar una ecuación para el costo esperado por unidad de tiempo, encontraremos los valores óptimos de n y k que minimicen este costo.[2]

Suponga que T, el número de periodos que el sistema permanece bajo control siguiendo a un ajuste, es una variable aleatoria discreta que tiene una distribución geométrica. Es decir,

$$P\{T = t\} = \pi(1 - \pi)^t \qquad \text{para } t = 0, 1, 2, 3, \ldots$$

---

[2] Se usaron ideas similares en los modelos de control de inventarios de los capítulos 4 y 5 y también se emplearán para analizar los modelos de remplazo por edad en la sección 12.7.

El modelo geométrico surge de la siguiente manera. Suponga que en cualquier periodo el proceso está bajo control. Entonces π es la probabilidad condicional de que el proceso se desvíe fuera de control en el siguiente periodo. La distribución geométrica es el análogo discreto de la distribución exponencial. Al igual que la distribución exponencial, la distribución geométrica también tiene la propiedad de amnesia.[3] En el presente contexto, la propiedad de amnesia implica que no hay envejecimiento ni decaimiento en el proceso de producción. Es decir, es igualmente posible que el proceso se desvíe fuera de control después que se ha encontrado una causa asignable y se ha corregido, como muchos periodos después. Esta hipótesis será exacta si las desviaciones del proceso se deben a causas aleatorias o si el proceso se recalibra sobre la marcha.

Una señal fuera de control se indica cuando

$$|\overline{X} - \mu| > \frac{k\sigma}{}.$$

Igual que en secciones anteriores de este capítulo, sea α la probabilidad del error Tipo 1. Este error ocurre cuando se observa una señal fuera de control aunque el proceso esté bajo control. Se concluye que

$$\alpha = P\left\{\left|\overline{X} - \mu\right| > \frac{k\sigma}{} \;\middle|\; E(\overline{X}) = \mu\right\}$$

$$= P\left\{\left|\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}\right| > k \;\middle|\; E(\overline{X}) = \mu\right\} = P\{|Z| > k\} = 2\Phi(-k),$$

donde Φ es la función acumulada de la distribución normal estándar.

La probabilidad del error Tipo 2, β, es la probabilidad de no detectar una condición fuera de control. Aquí suponemos que una condición fuera de control implica que la media del proceso se ha desviado a $\mu + \delta\sigma$ o $\mu - \delta\sigma$. Suponga que condicionamos al evento de que la media se ha desviado de μ a $\mu + \delta\sigma$. La probabilidad de que la desviación no se detecte después de observar una muestra de n observaciones es

$$\beta = P\left\{\left|\overline{X} - \mu\right| \leq \frac{k\sigma}{} \;\middle|\; E(\overline{X}) = \mu + \delta\sigma\right\}$$

$$= P\left\{\frac{-k\sigma}{\sqrt{n}} \leq \overline{X} - \mu \leq \frac{k\sigma}{} \;\middle|\; E(\overline{X}) = \mu + \delta\sigma\right\}$$

$$= P\left\{-k - \delta\sqrt{n} \leq \frac{\overline{X} - \mu - \delta\sigma}{\sigma/\sqrt{n}} \leq k - \delta\sqrt{n} \;\middle|\; E(\overline{X}) = \mu + \delta\sigma\right\}$$

$$= P\{-k - \delta\sqrt{n} \leq Z \leq k - \delta\sqrt{n}\}$$

$$= \Phi(k - \delta\sqrt{n}) - \Phi(-k - \delta\sqrt{n}).$$

Si hubiéramos condicionado a $E(\overline{X}) = \mu - \delta\sigma$, habríamos obtenido

$$\beta = \Phi(k + \delta\sqrt{n}) - \Phi(-k + \delta\sqrt{n}).$$

Usando la simetría de la distribución normal [específicamente, que $\Phi(t) = 1 - \Phi(-t)$ para cualquier t], es fácil demostrar que estas dos expresiones para β son las mismas.

[3] En la sección 12.2 se presenta un análisis detallado de la propiedad de amnesia de la distribución exponencial.

Considere las variables aleatorias T y S. Suponemos que T es una variable aleatoria geométrica que adopta valores 0, 1, 2, . . . Se puede demostrar que

$$E(T) = \frac{1 - \pi}{\pi}$$

(véase, por ejemplo, DeGroot, 1986). La variable aleatoria S es el número de periodos que el proceso permanece fuera de control después que ocurre una desviación. La probabilidad de que la desviación no se detecte cuando el proceso está fuera de control es exactamente β. Se concluye que S es también una variable aleatoria geométrica excepto que adopta solamente los valores 1, 2, 3, . . . Es decir,

$$P\{S = s\} = (1 - \beta)\beta^{s-1} \qquad \text{para } s = 1, 2, 3, \ldots$$

Como S se define sobre el conjunto 1, 2, 3, . . . , el valor esperado de S es $E(S) = 1/(1 - \beta)$. Se concluye que la longitud esperada de ciclo, digamos C, está dada por

$$E(C) = E(T + S) = E(T) + E(S) = \frac{1 - \pi}{\pi} + \frac{1}{1 - \beta}.$$

Considere el costo de muestreo esperado incurrido en un ciclo. En cada periodo hay n artículos muestreados. Así como hay en promedio $E(C)$ periodos por ciclo, se concluye que el costo de muestreo por ciclo es $a_1 nE(C)$.

Ahora calculamos el costo esperado de búsqueda. El proceso se detiene cada vez que se observa una señal de fuera de control. Una o más de estas señales podrían ser una falsa alarma. Suponga que hay exactamente M falsas alarmas en un ciclo. La variable aleatoria M tiene la distribución binomial con probabilidad de "éxito" (es decir, una falsa alarma) de α para un total de T pruebas. Se concluye que $E(M) = \alpha E(T)$. El número esperado de búsquedas por ciclo es exactamente $1 + E(M)$, pues se supone que la búsqueda final descubre y corrige la causa asignable. Así, el costo total de búsqueda en un ciclo es

$$a_2[1 + \alpha E(T)] = a_2[1 + \alpha(1 - \pi)/\pi].$$

También suponemos que hay un costo de a_3 para cada periodo que el proceso es operado en una condición fuera de control. El proceso está fuera de control para exactamente S periodos. Así, el costo esperado fuera de control es $a_3 E(S) = a_3/(1 - \beta)$.

Se concluye que el costo esperado por ciclo es

$$a_1 nE(C) + a_2[1 + \alpha(1 - \pi)/\pi] + a_3/(1 - \beta).$$

La división entre la longitud esperada del ciclo, $E(C)$, da el costo promedio por unidad de tiempo como

$$a_1 n + \frac{a_2\left[1 + \alpha\dfrac{(1 - \pi)}{\pi}\right] + \dfrac{a_3}{1 - \beta}}{\dfrac{1 - \pi}{\pi} + \dfrac{1}{1 - \beta}}$$

$$= a_1 n + \frac{a_2\left[1 + \alpha\dfrac{(1 - \pi)}{\pi}\right] + \dfrac{a_3}{1 - \beta}}{\dfrac{1 - \beta(1 - \pi)}{(1 - \beta)\pi}}$$

$$= a_1 n + \frac{a_2(1 - \beta)[\pi + \alpha(1 - \pi)] + a_3\pi}{1 - \beta(1 - \pi)}.$$

Escribiremos lo anterior como $G(n, k)$ para indicar que la optimización requiere la búsqueda de los mejores valores de n y k, donde $n = 1, 2, 3, \ldots$ y $k > 0$. Observe que α depende de k y β depende tanto de n como de k. El objetivo es encontrar los valores de n y k que minimicen a $G(n, k)$. Este es un problema complejo de optimización porque tanto α como β requieren la evaluación de la función de la distribución normal acumulada.

Ejemplo *11.5*

Considere el ejemplo 11.1 de Wonderdisk, introducido en la sección 11.2. A Howard Hamilton le gustaría diseñar una gráfica $\overline{X}$ de una manera económicamente óptima. Con base en su experiencia con el proceso y en un análisis de la historia pasada de fallas, decide que la distribución geométrica describe con exactitud los cambios en el estado del proceso.

Para usar el modelo descrito en esta sección, debe estimar los diferentes costos y los parámetros del sistema. El primero es el costo de muestreo. Aquí el muestreo necesita que se mida la longitud de un brazo de rastreo. Esto requiere mover el brazo a una posición diferente, montarlo en un bastidor especial para protegerlo, y medir la longitud con calibradores diseñados para ese propósito. El proceso toma aproximadamente 12 minutos del tiempo de un técnico. El técnico gana $15 por hora, de modo que el costo de muestreo es $15/5 = $3 por artículo muestreado.

El segundo costo que debe estimarse es el costo de búsqueda. Generalmente el tiempo ocupado en la búsqueda de una causa asignable de una señal fuera de control es aproximadamente de 30 minutos. Si no se descubre un problema en ese lapso de tiempo, generalmente se supone que la señal de fuera de control fue una falsa alarma. Los brazos generan una ganancia para la compañía de aproximadamente $1 200 al día. Suponiendo una jornada de trabajo de ocho horas, el costo de parar la producción es de aproximadamente $1 200/8 = $150 por hora. Así, el costo de búsqueda es de $75.

El tercer costo requerido por el modelo es el costo de operación del proceso en una condición fuera de control. Si el proceso está fuera de control, aumenta la proporción de brazos defectuosos producidos. La mayoría de los brazos defectuosos aparecen en la fase final de pruebas de las unidades de disco. Si una unidad de disco tiene un brazo defectuoso, la unidad de disco se desarma y el brazo se reemplaza. Algunos brazos defectuosos pasan la inspección y se embarcan al consumidor con la unidad de disco. Wonderdisk ofrece a los compradores una garantía de 14 meses, y es posible que se desarrolle un problema con la unidad de disco durante el periodo de la garantía si el brazo está defectuoso. Howard estima que el costo de operación del proceso fuera de control es aproximadamente $300 por hora, pero él no confía mucho en esta estimación.

El modelo también requiere estimaciones de π y δ. Recuerde que π representa la probabilidad de que el proceso se desvíe desde un estado bajo control hasta un estado fuera de control durante un periodo. En el pasado, las señales fuera de control han ocurrido con una tasa de aproximadamente una por cada 10 horas de operación. Como la mitad de éstas han sido falsas alarmas, una estimación razonable de la proporción de periodos en los cuales ha ocurrido una desviación es aproximadamente uno en cada 20, o sea $\pi = .05$. La constante δ representa el grado de la desviación medida en múltiplos de la desviación estándar del proceso. En el pasado, el promedio de las desviaciones ha sido aproximadamente de una desviación estándar, de modo que la estimación de δ es 1.

Para simplificar los cálculos, Howard decide usar una aproximación a la función de distribución acumulada normal estándar. La que usa es la siguiente:

$$\Phi(z) = 0.500232 - 0.212159z^{2.08388} + 0.5170198z^{1.068529} + 0.041111z^{2.82894}.$$

Esta aproximación, debida a Herron (1985), tiene una exactitud de 0.5 por ciento para $0 < z < 3$. Howard Hamilton decide que esto es suficientemente exacto para sus propósitos. El esquema de optimización que adopta es el siguiente. Como n es una variable discreta que representa el número de artículos muestreados en cada subgrupo, es poco probable que n sobrepase a 10. Además, puesto que k es el número de las desviaciones estándar de $\overline{X}$ que se usan en la gráfica de control, es poco probable que k sobrepase a 3. Howard escribe un programa de computadora para evaluar $G(n, k)$ para $k = 0, 0.1, 0.2, \ldots, 2.8, 2.9, 3.0$ y $n = 1, 2, 3, \ldots, 10$. (En realidad estos cálculos se hicieron usando un popular programa de hoja de cálculo.) Para cada valor fijo de n, la función $G(n, k)$ resulta ser convexa para la variable k (las funciones convexas se estudiaron en los capítulos 4 y 5). Para los valores paramétricos dados y para $n = 4$, la figura 11-12 muestra a la función $G(n, k)$ como una función de k. La gráfica muestra que el costo mínimo ocurre para más o menos $k = 1.7$ y es igual a aproximadamente $45 por hora.

FIGURA 11-12
El comportamiento de $G(n, k)$ como una función de k (refiérase al ejemplo 11.5).

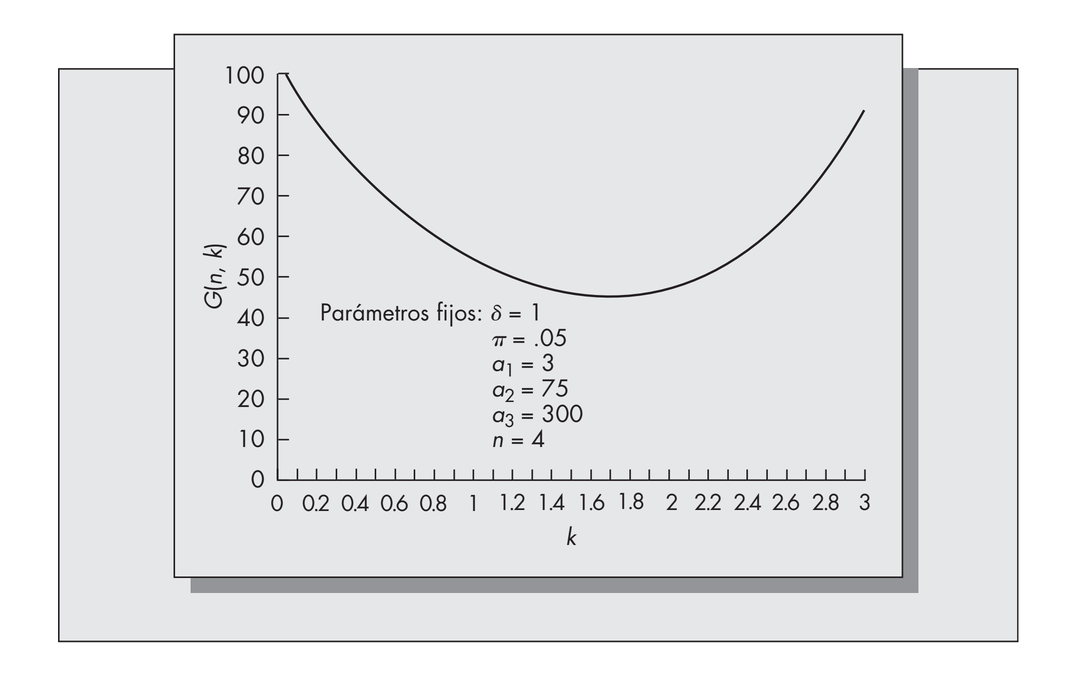


TABLA 11-4 **Valores óptimos de *k* para diferentes valores de *n***

Los parámetros fijos son: $\delta = 1$, $\pi = 0.05$, $a_1 = 3.$, $a_2 = 75$, $a_3 = 300.$

n	k óptima	α	β	Costo
1	1.13	.25	.54	$54.4
2	1.52	.14	.50	48.8
3	1.60	.11	.44	46.1
4	1.74	.08	.39	45.2
5	1.86	.07	.34	45.3
6	1.97	.05	.30	46.2
7	2.06	.04	.26	47.6
8	2.14	.03	.23	49.3
9	2.21	.02	.21	51.3
10	2.27	.02	.18	53.4

La tabla 11-4 da los resultados de los cálculos para estas asignaciones de parámetros. De acuerdo con la tabla, el tamaño óptimo del subgrupo es 4 y la k óptima es 1.74. Estos resultados se acercaron razonablemente a la política actual y le dan confianza a Hamilton de que cuando menos sus estimaciones estuvieron del lado correcto. Sin embargo, cuando él presentó los resultados a su jefe, éste expresó preocupación en torno al grande valor de β, la probabilidad del error Tipo 2:

> ¿Quiere usted decir que si operamos el sistema en forma óptima, entonces hay casi un 40 por ciento de posibilidad de que no podremos detectar cuando el sistema se ha salido de control? Eso suena bastante alto. Considerando el empuje que existe en toda la compañía por una mejora de la calidad, yo pienso que esta cifra es muy inquietante. ¿Cómo es que eso resulta ser óptimo?

La razón de que el error Tipo 2 fuera tan grande era el supuesto de que el costo para la compañía fuera de $300 por cada hora que el proceso se operara fuera de control. Después de pensar un poco

TABLA 11-5 Valores óptimos de k para diferentes valores de n (revisado). Parámetros fijos: los mismos que en la tabla 11-3 excepto $a_3 = 1\,000$

n	k óptima	α	β	Costo
1	0.67	.48	.30	$111.6
2	0.93	.33	.29	102.8
3	1.13	.25	.26	96.9
4	1.35	.18	.25	93.3
5	1.43	.15	.21	91.3
6	1.59	.11	.19	90.5
7	1.70	.09	.16	90.4
8	1.79	.07	.14	90.9
9	1.86	.06	.12	91.8
10	1.94	.05	.11	93.3

sobre el asunto, el jefe de Howard decidió que un valor de $1 000 era probablemente más cercano a la marca y era más consistente con la meta de la corporación de mejorar la calidad. Howard repitió sus cálculos sustituyendo un valor de $a_3 = 1\,000$. Los resultados se presentan en la tabla 11-5.

Con el valor revisado de $a_3 = 1\,000$, el valor óptimo del tamaño de la muestra n aumentó a 7 con un valor correspondiente de $k = 1.70$. El costo por hora para la solución óptima aumentó a $90.40. Aunque los límites en la gráfica de control permanecieron iguales, aumentar el tamaño de la muestra de 4 a 7 condujo a una dramática disminución del valor de β de .39 a .16. El nuevo jefe de Howard estuvo mucho más feliz con los valores de α y de β que resultaron del nuevo diseño.

Problemas para la sección 11.6

23. Un ingeniero de control de calidad está considerando el diseño óptimo de una gráfica $\overline{X}$. Con base en su experiencia con el proceso de producción, existe una probabilidad de .03 de que el proceso se desvíe de un estado bajo control a otro fuera de control en cualquier periodo. Si el proceso se desvía fuera de control, puede atribuirse a una causa asignable individual; la magnitud de la desviación es 2σ. Las muestras de n artículos se hacen en forma horaria, y cada muestreo cuesta $0.50 por unidad. Buscar la causa asignable cuesta $25, y operar el proceso bajo un estado fuera de control cuesta $300 por hora.

    a) Determine el costo horario de operar el sistema si $n = 6$ y $k = 2.5$.

    b) Estime el valor óptimo de k para el caso de $n = 6$. Si está haciendo sus cálculos a mano, use $k = 0.5, 1, 1.5, 2, 2.5$ y 3.0. Si usted está empleando una computadora, use $k = 0.1, 0.2, \ldots, 2.9, 3.0$.

    c) Determine el diseño óptimo de la gráfica de control que minimice los costos anuales promedio.

24. Considere la aplicación del diseño económico de las gráficas $\overline{X}$ para Wonderdisk presentado en esta sección. Sin ejecutar realmente los cálculos, discuta cuál sea posiblemente el efecto sobre los valores óptimos de n y k si

    a) δ aumenta de 1 a 2.

    b) π aumenta de .05 a .10.

    c) a_1 disminuye a 1.

    d) a_2 aumenta a 150.

25. ¿Bajo qué circunstancias no serían exactos los siguientes supuestos?

   a) El supuesto de que la ley de probabilidad que describe al número de periodos hasta que el proceso se sale de control sigue la distribución geométrica.

   b) El supuesto de que una condición fuera de control corresponde a una desviación de la media igual a $\delta\sigma$.

   c) El supuesto de que el costo de búsqueda es una constante fija, a_2.

26. Discuta las siguientes posiciones en pro y en contra sobre el uso de modelos de optimización para diseñar las gráficas de control:

   Contra: "Estos modelos son inútiles para mí porque no siento que pueda estimar con exactitud los valores de los datos requeridos".

   Pro: "La selección de valores específicos de n y k en la construcción de las gráficas X implica que usted está suponiendo valores para los diferentes costos y parámetros del sistema. De igual forma, usted podría enfrentar el problema y obtener las mejores estimaciones que pueda y usarlas para diseñar la gráfica X."

*27. Suponga que "la salida del proceso fuera de control" corresponde a una desviación de la media de μ a $\mu + \sigma$ con probabilidad .25, $\mu + 2\sigma$ con probabilidad .25, $\mu - \sigma$ con probabilidad .25, y $\mu - 2\sigma$ con probabilidad .25. ¿Qué modificaciones se requieren en el modelo? En especial, si expresamos a la desviación en la forma $\mu \pm \delta\sigma$, muestre cómo calcular β_1 y β_2 que corresponderían a los valores de $\delta = 1$ y $\delta = 2$, respectivamente. Si ahora los costos de operar fuera de control estuvieran representados por a_3 cuando $\delta = 1$ y a_4 cuando $\delta = 2$, determine una ecuación para los costos de operación promedio anuales. (Suponga que todos los otros costos y los parámetros del sistema permanecen iguales.)

28. Un contratista local fabrica las campanas que se usan en los teléfonos. La compañía telefónica requiere que las campanas suenen para un nivel específico de ruido (en decibeles). Se está diseñando una gráfica $\overline{X}$ para monitorear esta variable. El proceso de muestreo de las campanas provenientes de la línea requiere golpear a las campanas con un disco basculante de fuerza fija y medir el nivel de decibeles en un medidor diseñado para ese propósito. El costo del muestreo es de $1.25 por campana. Si proceso se sale de control, el espesor de las campanas es incorrecto. El costo de buscar una causa asignable se estima en $50. El costo de operar el proceso en un estado fuera de control se estima en $180 por hora. La operación fuera de control corresponde a una desviación de 2σ en el nivel de decibeles, y la probabilidad de que el proceso se desvíe fuera de control en cualquier hora es .03.

   a) La compañía usa una gráfica $\overline{X}$ basada en límites 3σ y subgrupos de tamaño 4. ¿Cuál es el costo horario?

   b) ¿Cuáles son los valores óptimos de n y k para este proceso y los costos óptimos asociados?

11.7 PANORAMA GENERAL DE MUESTREO DE ACEPTACIÓN

Las gráficas de control proporcionan una forma cómoda de monitorear un proceso en tiempo real para determinar si ha ocurrido una desviación en los parámetros del proceso. Otro aspecto importante del control de calidad es determinar la calidad de los bienes manufacturados *después* que han sido producidos. En la mayoría de los casos la inspección al 100 por ciento es imposible o impráctica. Así, se inspecciona una muestra de artículos y se estiman los parámetros de calidad de lotes grandes de artículos con base en los resultados del muestreo.

Para ser más específicos, el muestreo de aceptación encara el siguiente problema: si se extrae una muestra de un lote grande de artículos y la muestra está sujeta a una inspección al 100

por ciento, ¿qué inferencias pueden obtenerse acerca de la calidad del lote con base en la calidad de la muestra? El análisis estadístico proporciona un medio para extrapolar las características de una muestra a las características del lote, y un medio para determinar la probabilidad de llegar a la conclusión equivocada.

Obviamente, la inspección al 100 por ciento de todos los artículos del lote reducirá a cero la probabilidad de una conclusión incorrecta. Sin embargo, hay varias razones por las cuales la inspección al 100 por ciento no es factible o no es deseable. Algunas de éstas incluyen

1. En la mayoría de los casos la inspección al 100 por ciento es demasiado costosa. Es virtualmente imposible para las líneas de transferencia de alto volumen y para los procesos continuos de producción.
2. En algunos casos la inspección al 100 por ciento puede ser imposible, como cuando la inspección incluye la prueba destructiva del artículo. Por ejemplo, determinar la vida útil de un foco requiere mantenerlo encendido hasta que se funda.
3. Si la inspección la hace el consumidor en lugar del productor, la inspección al 100 por ciento del consumidor proporciona poco incentivo al productor para que mejore la calidad. Es más barato para el productor reparar o reemplazar los artículos devueltos por el consumidor que mejorar la calidad del proceso de producción. Sin embargo, si el consumidor devuelve el lote completo con base en los resultados del muestreo, esto proporciona una motivación mucho mayor al productor para que mejore la calidad de los siguientes lotes.

En este capítulo tratamos los tres siguientes planes de muestreo.

1. *Planes de muestreo simple*. Los planes de muestreo simple son por mucho los más populares y los más fáciles de usar de los planes que estudiaremos. Dos números, n y c, determinan un plan de muestreo simple. Si hay más de c artículos defectuosos en una muestra de tamaño n, se rechaza al lote; de otra manera se acepta.

2. *Planes de muestreo doble*. En un plan de muestreo doble, primero seleccionamos una muestra de tamaño n_1. Si el número de artículos defectuosos en la muestra es menor o igual que c_1, se acepta el lote. Si el número de artículos defectuosos en la muestra es mayor que c_2, entonces se rechaza. Sin embargo, si el número de artículos defectuosos es mayor que c_1 y menor o igual que c_2, se extrae una segunda muestra de tamaño n_2. Ahora se acepta el lote si el número acumulado de artículos defectuosos en ambas muestras es menor o igual a un tercer número, c_3. (Frecuentemente $c_3 = c_2$.)

3. *Muestreo secuencial*. Es obvio que un plan de muestreo doble puede ampliarse a un plan de muestreo triple, el cual puede ampliarse a un plan de muestreo cuádruple, etcétera. Un plan de muestreo secuencial es la conclusión lógica de este proceso. Los artículos se muestrean uno a la vez y se registra el número acumulado de artículos defectuosos para cada etapa del proceso. Con base en el valor del número acumulado de artículos defectuosos, hay tres decisiones posibles en cada etapa:

a) Rechazar el lote.

b) Aceptar el lote.

c) Continuar el muestreo.

Un plan de muestreo complejo puede tener propiedades estadísticas deseables, pero podría ser difícil calcular las regiones de aceptación y rechazo y el plan podría ser difícil de implementar. El plan de muestreo correcto para un ambiente específico puede no ser el de mayor sofisticación matemática. Al igual que con cualquier herramienta analítica, deben sopesarse los beneficios potenciales contra los costos potenciales.

Kolesar (1993) puntualiza que al mejorar los estándares de calidad, el valor del muestreo de aceptación puede disminuir en los años venideros. Motorola se ha hecho famoso por instituir su empuje de calidad "seis sigma" (las iniciativas de calidad de Motorola se discuten en la Aplicación Instantánea de la página 660). Con esto, Motorola quiere decir que la tasa de

defectos no debe ser mayor que el área fuera de $\pm 6\sigma$ bajo la curva normal. Esto se traduce en tasas de defectos menores que 3.4 partes por millón. (Pocas tablas de la distribución normal van más allá de 4σ, así es que va a ser un poco difícil verificar esta probabilidad.) Si las tasas de defectos son muy bajas, el muestreo de aceptación se hace muy ineficiente. Por ejemplo, suponga que la tasa de defectos aumentara por un factor de 10. En ese caso, la probabilidad de encontrar un defecto de, digamos, 1 000 unidades sería solamente de .034. No esperaríamos ver un solo defecto hasta que hubiéramos muestreado cuando menos 29 lotes en promedio. Sin embargo, debemos tener en mente que el estándar de calidad seis sigma de Motorola no se ha convertido en un estándar industrial ni por asomo. El muestreo de aceptación seguirá siendo una herramienta valiosa durante muchos años venideros.

Aplicación instantánea

NAVISTAR SE ANOTA PUNTOS CON EL PROGRAMA DE CALIDAD SEIS-SIGMA

Navistar International es un fabricante estadounidense importante de camiones, autobuses, y motores y tiene varias plantas alrededor del mundo. En 1985 la fuerza de trabajo mundial de Navistar sobrepasaba los 110 000. Debido a una destructora huelga de de la unión de trabajadores de la industria automotriz (UAW, por United Auto Workers) y a una recesión, la compañía tuvo que recortar severamente la fuerza laboral para sobrevivir. En la actualidad la fuerza laboral es aproximadamente de 20 000. Para combatir los problemas de costo y calidad que estaba experimentando en ese momento, Navistar decidió lanzar un programa de calidad seis sigma a mediados de la década de los 90. Como se observa en esta sección, seis sigma implica tasas de defectos de 3.4 partes por millón o menores. Aún cuando los programas seis sigma rara vez alcanzan tasas de defectos tan bajas, la meta es clara: haga lo que sea necesario en la organización para realizar un cambio fundamental en las actitudes tanto de la gerencia como de los trabajadores acerca de la calidad. Sin embargo, los programas de calidad no son gratis. Navistar pagó a una compañía consultora más de $6 millones para implementar este programa. Un resultado inmediato fue que el precio de las acciones de Navistar aumentó más del 400 por ciento en los 14 meses siguientes a la implementación del programa. (Por supuesto, como todos sabemos, el precio de la acción de una compañía está influido por muchos factores, así que no está claro cuál es el papel que jugó el programa de seis sigma.)

Los programas seis sigma tienen su propia cultura. A los empleados especialmente entrenados se les otorgan cintas negras después un mes de entrenamiento, y cintas negras de maestros después de entrenamiento adicional. A los cintas negras se les asignan proyectos específicos y tienen el poder de ir directamente a la gerencia general con soluciones propuestas. Por supuesto, para que funcione este enfoque, no solamente los empleados, sino también la gerencia, deben comprometerse firmemente con el programa. ¿Cree todo el mundo en el valor de los programas seis sigma? Evidentemente no; por ejemplo, Charles Holland, presidente de una compañía consultora con base en Knoxville que se especializa en los métodos del control estadístico de la calidad, considera al programa seis sigma como una “bala de plata” vendida a “precios escandalosos”.[1]

Si esto es verdad, ¿qué motivó a Navistar a soltar $6 millones para este programa? De acuerdo con John Horne, el ejecutivo en jefe de la compañía en 1995, la compañía necesitaba un antídoto para la decadencia que estaba experimentando: “No teníamos una estrategia; la mayoría de las compañías no la tienen”. La estrategia que Horne adoptó fue perseguir los problemas de la compañía al nivel de la planta. Los problemas de control de calidad habían estado evadiendo a las plantas de Navistar durante años. El objetivo del programa seis sigma era la masiva planta de 4 000 pies cuadrados en Springfield, Ohio. (Navistar no implementó el seis sigma en todas sus plantas por varias razones. Por ejemplo, la oposición del sindicato evitó la implementación en la planta del Canadá ubicada en Chatham, Ontario.)

¿Cuál fue el resultado en Springfield? El esfuerzo ha sido recompensando con $1 millón en ahorros el primer año, y mayores ahorros en años subsiguientes. Se proyectó que el ahorro total solamente en esta planta sería de $26 millones, muy arriba del costo de $6 millones del programa. Algunas veces el kaizen (mejora continua) simplemente no es suficiente para arreglar un sistema en problemas. Aun cuando es caro, seis sigma puede proveer el empujón inicial necesario para hacer que las cosas cambien, como lo hizo con Navistar.

[1] Franklin, S., “In Pursuit of Perfection”, *Chicago Tribune*, domingo 4 de abril, 1999, sección 5, pp. 7-8.

11.8 NOTACIÓN

Usaremos la siguiente notación en lo que resta de este capítulo.

N = número de piezas en un lote o un grupo dados.

n = Número de piezas en la muestra ($n < N$).

M = Número de artículos defectuosos en el lote.

β = Riesgo del consumidor: la probabilidad de aceptar lotes malos.

α = Riesgo del productor: la probabilidad de rechazar lotes buenos.

c = Nivel de rechazo.

X = Número de artículos defectuosos en la muestra.

p = Proporción de artículos defectuosos en el lote.

p_0 = Nivel aceptable de calidad (AQL, *acceptable quality level*).

p_1 = Tolerancia porcentual de artículos defectuosos del lote (LTPD, *lot tolerance percent defective*).

Suponga que N es una constante conocida. Si N es muy grande en relación con el tamaño de la muestra, n, puede suponerse como infinita. En ese caso no entra en los cálculos. Aunque M también es una constante, su valor *no* se conoce anticipadamente. De hecho, sólo la inspección al 100 por ciento revelará el valor verdadero de M. Con frecuencia nos interesa el análisis del comportamiento del plan de muestreo para diferentes valores de M. El riesgo del consumidor y el riesgo del productor dependen del plan de muestreo. Finalmente, X, el número de artículos defectuosos en la muestra, es una variable aleatoria. Esto significa que si tuviéramos que repetir el experimento de muestreo con una muestra aleatoria diferente de tamaño n, no observaríamos necesariamente el mismo número de artículos defectuosos. Con base en las propiedades estadísticas de la población como un todo, podemos determinar la forma de la distribución de probabilidad de X.

El nivel aceptable de calidad, p_0, es el nivel deseado o meta de la proporción de artículos defectuosos en el lote. Si la proporción verdadera de artículos defectuosos en el lote es menor o igual a p_0, se considera que el lote es aceptable. La tolerancia porcentual de artículos defectuosos, p_1, es una proporción inaceptable de artículos defectuosos en el lote. Se considera que el lote es inaceptable si la proporción de artículos defectuosos sobrepasa a p_1. Debido a la imprecisión del muestreo estadístico, permitimos un área gris entre p_0 y p_1. Si AQL y LTPD son iguales, pueden requerirse tamaños grandes de muestra para alcanzar valores aceptables de α y β.

11.9 MUESTREO ÚNICO PARA ATRIBUTOS

El objetivo de todos los procedimientos de muestreo consiste en estimar las propiedades de una población a partir de las propiedades de la muestra. En especial, deseamos probar las hipótesis

$$H_0\text{: El lote tiene una calidad aceptable } (p \le p_0).$$

$$H_1\text{: El lote tiene una calidad inaceptable } (p \ge p_1).$$

La prueba es de la forma: rechace H_0 si $X > c$. El valor de c depende de la selección de β, la probabilidad del error Tipo 1. La probabilidad del error Tipo 1 es la probabilidad de rechazar H_0 cuando es verdadera. En el contexto del problema del control de la calidad, ésta es la probabilidad de rechazar el lote cuando es aceptable. Esto también se conoce como el riesgo del productor. En forma de ecuación,

$$\begin{aligned}\alpha &= P\{\text{Rechazar } H_0 \mid H_0 \text{ es verdadera}\} = P\{\text{Rechazar el lote} \mid \text{El lote es bueno}\}\\ &= P\{X > c \mid p = p_0\}.\end{aligned}$$

La distribución exacta de X es *hipergeométrica* con parámetros n, N y M. Es decir,

$$P\{X = m\} = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}} \qquad \text{para } 0 \le m \le \text{mín}(M, n),$$

donde

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}.$$

En la mayoría de las aplicaciones, N es mucho mayor que n, de modo que es satisfactoria la aproximación binomial a la hipergeométrica. En ese caso

$$P\{X = m\} = \binom{n}{m} p^m (1-p)^{n-m} \qquad \text{para } 0 \le m \le n,$$

donde $p = M/N$ es la proporción verdadera de artículos defectuosos en el lote.

Usando la aproximación binomial, el riesgo del productor y el riesgo del consumidor están dados por

$$\alpha = P\{X > c \mid p = p_0\} = \sum_{m=c+1}^{n} \binom{n}{m} p_0^m (1-p_0)^{n-m},$$

$$\beta = P\{X \le c \mid p = p_1\} = \sum_{m=0}^{c} \binom{n}{m} p_1^m (1-p_1)^{n-m}.$$

La mayoría de las pruebas estadísticas requieren la especificación de la probabilidad del error Tipo 1, α. Los valores de α, n y p_0 determinarán un valor único de c, que puede obtenerse en tablas de la distribución binomial acumulada. Sin embargo, como la binomial es una distribución discreta, tal vez no sea posible encontrar c que concuerde exactamente con el valor deseado de α. Si p es pequeño y n es moderadamente grande ($n > 25$ y $np < 5$), la distribución de Poisson proporciona una aproximación adecuada a la binomial. Para valores muy grandes de n tales que $np(1-p) > 5$, la distribución normal provee una aproximación adecuada a la binomial. Consulte el apéndice 11-A para un análisis detallado de estas aproximaciones.

Ejemplo *11.6*

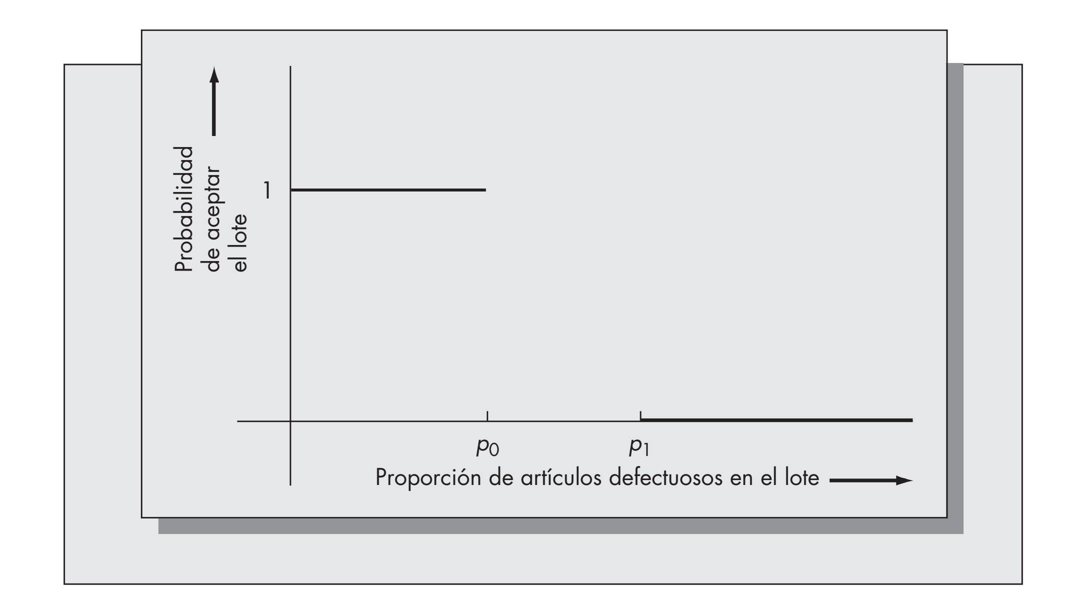
Spire Records es una gran cadena de tiendas al menudeo de la Costa Oeste que se especializa en discos y cintas. Uno de los proveedores de Spire es B&G Records, que envía discos a Spire en lotes de 100 discos. Después de alguna negociación, Spire y B&G han acordado que una tasa de 10 por ciento de artículos defectuosos es aceptable y una tasa de 30 por ciento de artículos defectuosos es inaceptable. De cada lote de 100 discos, Spire ha establecido el siguiente plan de muestreo: 10 discos se muestrean, y si se encuentra que más de 2 están torcidos, rayados, o son defectuosos de alguna otra manera, se rechaza el lote. Considere el riesgo del consumidor y del productor asociado con este plan de muestreo.

A partir de la información dada, tenemos que $p_0 = .1$, $p_1 = .3$, $n = 10$, y $c = 2$. Entonces,

$$\alpha = P\{X > c \mid p = p_0\} = P\{X > 2 \mid p = .1\} = 1 - P\{X \le 2 \mid p = .1\}$$
$$= 1 - \sum_{k=0}^{2} \binom{10}{k} (.1)^k (.9)^{10-k} = 1 - .9298 = .0702.$$

$$\beta = P\{X \le c \mid p = p_1\} = P\{X \le 2 \mid p = .3\}$$
$$= \sum_{k=0}^{2} \binom{10}{k} (.3)^k (.7)^{10-k} = .3828.$$

FIGURA 11-13
La curva OC ideal.



Observe que los valores de los parámetros $n = 10$, $p = .1$ y $n = 10$, $p = .3$ implican que ni la distribución normal ni la aproximación de Poisson son exactas, (El lector puede verificar con la tabla A-3 al final de este libro, que el uso de la distribución de Poisson con $\lambda = np$ proporciona valores aproximados para α y β, los cuales son .0803 y .4216, respectivamente.)

Derivación de la curva OC

La curva característica de operación (OC, de *operating characteristic*) mide la efectividad de una prueba para separar lotes de calidad variable. La curva OC es una función de p, la proporción verdadera de artículos defectuosos en el lote, y está dada por

$$\text{OC}(p) = P\{\text{Aceptar el lote} \mid \text{Proporción verdadera de artículos defectuosos} = p\}.$$

Ahora derivaremos la forma de la curva OC para el caso particular de un plan de muestreo simple con una muestra de tamaño n y nivel de rechazo c. En ese caso,

$$\begin{aligned}\text{OC}(p) &= P\{X \le c \mid \text{Proporción de artículos defectuosos en el lote} = p\} \\ &= \sum_{k=0}^{c} \binom{n}{k} p^k (1-p)^{n-k}.\end{aligned}$$

Lo ideal es que el procedimiento de muestreo pueda distinguir perfectamente entre los lotes buenos y los malos. La figura 11-13 muestra la curva OC ideal.

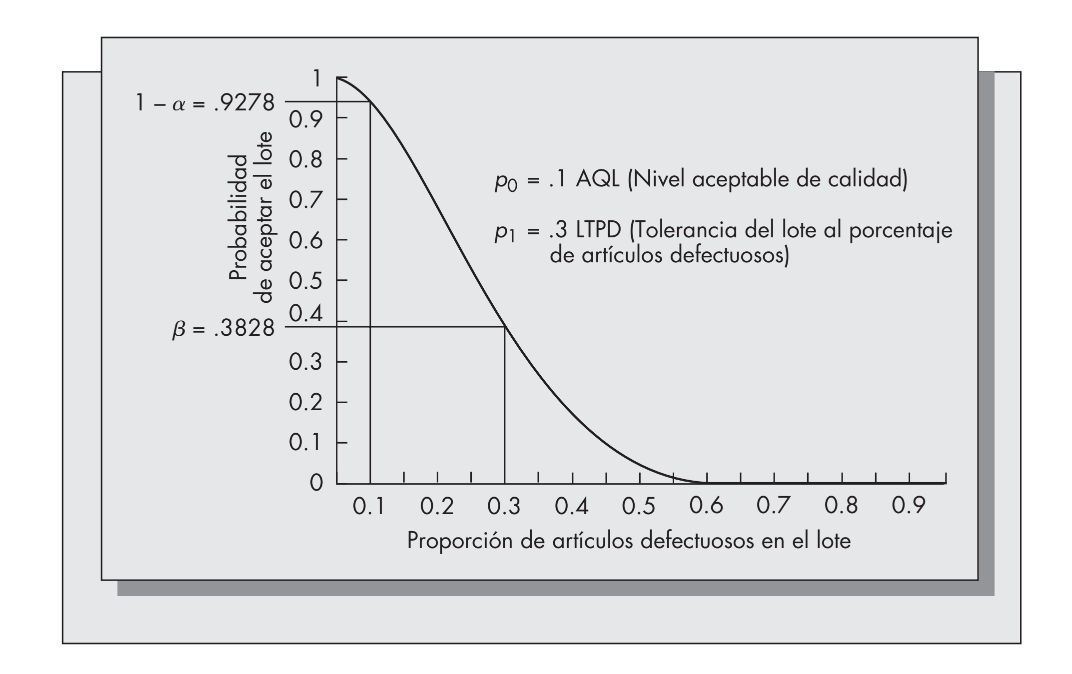
Ejemplo *11.6* *(continuación)*

Considere nuevamente el ejemplo 11.6 de Spire Records. La curva OC para su plan de muestreo simple está dada por

$$\text{OC}(p) = \sum_{k=0}^{2} \binom{10}{k} p^k (1-p)^{10-k}.$$

La gráfica de la curva OC de Spire se muestra en la figura 11-14. Un examen de la figura muestra que este plan de muestreo específico tiene más ventajas para el proveedor, B&G, que para Spire. El valor de $\beta = .3828$ implica que Spire está pasando casi 40 por ciento de los lotes que contienen 30 por ciento de artículos defectuosos. Además, no se descarta la probabilidad de aceptar lotes con proporciones de artículos defectuosos tan altos como el 40 por ciento y aún 50 por ciento. Esto toma en cuenta la experiencia de Spire de que parecía haber muchas devoluciones de los clientes de discos de marca B&G.

FIGURA 11-14
La curva OC para Spire Records ($n = 10$).



A Herman Sondle, un empleado de Spire inscrito en un programa de maestría local, se le pidió que examinara el problema de los discos B&G. Él descubrió la causa del problema analizando la curva OC ilustrada en la figura 11-14. Para disminuir las posibilidades de que Spire reciba lotes malos de B&G, sugirió que se modificara el plan de muestreo al hacer $c = 0$. El riesgo del consumidor resultante es

$$\beta = P\{X \leq 0 \mid p = .3\} = (.3)^0(.7)^{10} = .028,$$

o aproximadamente 3 por ciento. Esto parecía ser un nivel aceptable de riesgo, de modo que la compañía instituyó esta política. Desafortunadamente, la proporción de grupos rechazados aumentó dramáticamente. El valor resultante del riesgo del productor, α, es

$$\alpha = P\{X > 0 \mid p = .1\} = 1 - P\{X = 0 \mid p = .1\} = 1 - (.9)^{10} = .6513.$$

Es decir, Spire estaba rechazando aproximadamente 65% de grupos buenos. B&G amenazó con interrumpir los embarques a Spire a menos que regresara a su plan de muestreo original.

La gerencia de Spire no sabía que hacer. Si regresaba al plan original, enfrentaba el riesgo de perder clientes que irían a otro lado a comprar discos con más calidad. Si continuaba con el plan presente, se arriesgaba a perder a B&G como proveedor. Afortunadamente Sondle, quien había estado estudiando los métodos de control de calidad, pudo proponer una solución. Si se aumentara el tamaño de la muestra, la potencia de la prueba podría mejorar. Finalmente, podría diseñarse una prueba que tuviera niveles aceptables tanto para el riesgo del consumidor como para el del productor. Ya que B&G insistía en no más de una probabilidad del 10 por ciento de rechazar los lotes buenos, Spire también quería no más de una probabilidad del 10 por ciento de aceptar los lotes malos.

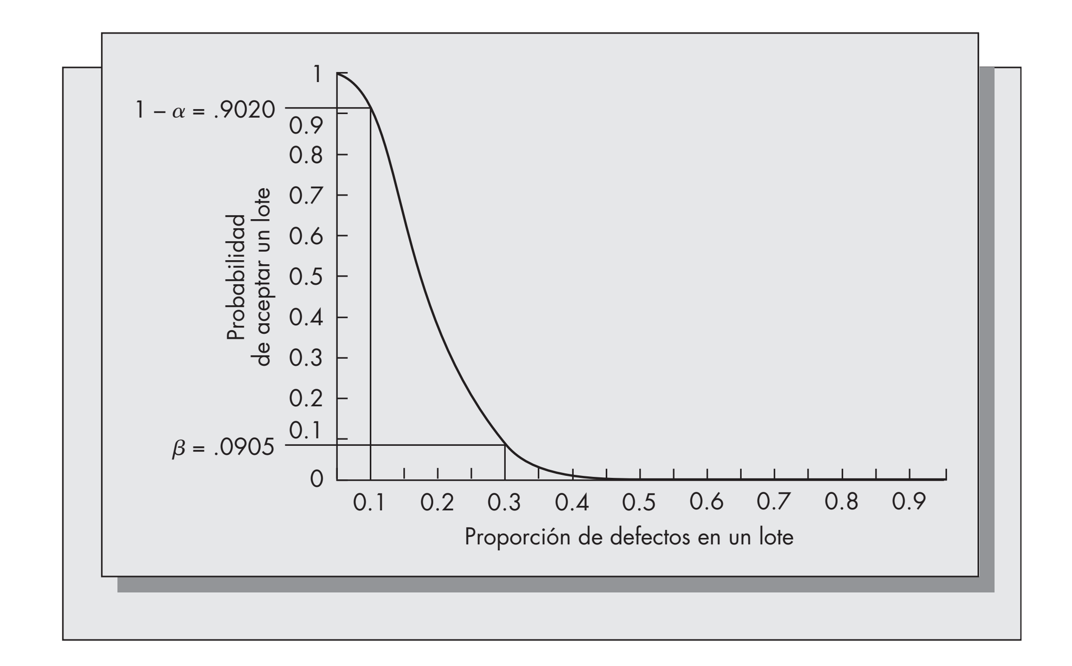
Después de algunos experimentos, Sondle encontró que un tamaño de muestra de $n = 25$ con un nivel de rechazo de $c = 4$ parecía cumplir con los requerimientos tanto de B&G como de Spire. Los valores exactos de α y β para esta prueba son

$$\alpha = P\{X > 4 \mid p = .1, n = 25\} = .0980.$$

$$\beta = P\{X \leq 4 \mid p = .3, n = 25\} = .0905.$$

Por supuesto que la eficiencia mejorada de este plan no sería gratis. El tiempo de empleado requerido para inspeccionar los discos de B&G se incrementó en dos y media veces. B&G y Spire estuvieron de acuerdo en compartir el costo adicional de la inspección. La curva OC del plan de muestreo con $n = 25$

FIGURA 11-15
Curva OC revisada para Spire Records ($n = 25$).



y c = 4 se muestra en la figura 11-15. Observe que esto se aproxima mucho más a la curva ideal que la curva OC para el plan original ilustrado en la figura 11-14.

Problemas para la sección 11.9

29. Se extraen muestras de tamaño 20 de lotes de 100 artículos y los lotes se rechazan si el número de artículos defectuosos en la muestra sobrepasa a 2. Si la proporción verdadera de artículos defectuosos en el lote es 5 por ciento, determine la probabilidad de aceptar un lote usando

    a) La distribución hipergeométrica exacta.

    b) La aproximación binomial a la hipergeométrica.

    c) La aproximación de Poisson a la binomial.

    d) La aproximación normal a la binomial.

30. Un productor de calculadoras de bolsillo compra los chips del procesador central en lotes de 1 000. Le gustaría tener una tasa de artículos defectuosos de 1 por ciento pero normalmente no va a rechazar un lote a menos que tenga 4 por ciento o más de artículos defectuosos. Se extraen muestras de 50 de cada lote, y el lote se rechaza si se encuentran más de dos artículos defectuosos.

    a) ¿Cuanto valen p_0, p_1, n, y c para este problema?

    b) Calcule α y β. Use la aproximación de Poisson para sus cálculos.

31. Una compañía emplea el siguiente plan de muestreo: extrae una muestra del 10 por ciento del lote que se está inspeccionando. Si 1 por ciento o menos de la muestra es defectuoso, se acepta el lote. De otra manera el lote se rechaza.

    a) Si un lote contiene 500 artículos de los cuales 10 son defectuosos, ¿cual es la probabilidad de que se acepte el lote?

b) Si un lote contiene 1 000 artículos de los cuales 20 son defectuosos, ¿cuál es la probabilidad de que se acepte el lote?

c) Si un lote contiene 10 000 artículos de los cuales 200 son defectuosos, ¿cual es la probabilidad de que se acepte el lote?

32. Hemispherical Conductor produce el microprocesador 80J84, el cual planea usar la compañía Sayle en una máquina cardiopulmonar. Debido a la sensibilidad de la aplicación, Sayle ha establecido un valor de AQL de .001 y de LTPD de .005. Sayle compra los microprocesadores en lotes de 500 y prueba 100 de cada lote. La prueba requiere destruir el microprocesador. El lote se rechaza si se encuentran artículos defectuosos.

    a) ¿Cuáles son los valores de p_0, p_1, n, y c que se usan?

    b) Calcule α y β.

    c) En vista de su respuesta al inciso *b*, ¿cuál pudo haber sido el problema de Sayle?

33. Determine un plan de muestreo para Spire Records que conduzca a $\alpha = .05$, $\beta = .05$, AQL = .10, y LTPD = .30. Discuta las ventajas y las desventajas del plan que usted obtenga en comparación con el plan actual de $n = 25$ y $c = 4$. (Consulte el ejemplo 11.6.)

*11.10 PLANES DE DOBLE MUESTREO PARA ATRIBUTOS

Cinco números definen a un plan de muestreo doble: n_1, n_2, c_1, c_2, y c_3. El plan se implementa de la siguiente manera: se extrae una muestra inicial de tamaño n_1 y se determina el número de artículos defectuosos en la muestra. Si el número de artículos defectuosos en la muestra es menor que o igual a c_1, se acepta el lote. Si el número de artículos defectuosos en la muestra es mayor que c_2, se rechaza el lote. Sin embargo, si el número de artículos defectuosos es mayor que c_1 pero menor que o igual que c_2, se extrae otra muestra de tamaño n_2. Si el número de artículos defectuosos en las muestras combinadas es menor que o igual que c_3, se acepta el lote. Si no, se rechaza el lote. La mayoría de los planes de muestreo doble suponen que $c_3 = c_2$. También formularemos esa hipótesis a partir de este momento.

Un plan de muestreo doble es obviamente más difícil de construir y más difícil de implementar que un plan de muestreo simple. Sin embargo, ciertamente tiene algunas ventajas sobre los planes simples. Primero, un plan de muestreo doble puede dar niveles similares de los riesgos del consumidor y del productor pero requiere menos muestreo a largo plazo que un plan simple. También existe la ventaja psicológica de los planes de muestreo doble de dar una segunda oportunidad a un lote antes de rechazarlo.

Ejemplo *11.7* Considere nuevamente el ejemplo 11.6 relacionado con Spire Records. Herman Sondle decide experimentar con algunos planes de muestreo doble para ver si pueden alcanzar niveles similares de eficiencia con menos muestreo. Desafortunadamente, como el plan dependerá de cuatro números diferentes, se necesita una considerable experimentación de prueba y error. Consideremos el cálculo de los riesgos del consumidor y del productor para el siguiente plan de muestreo:

$$n_1 = 20, \qquad c_1 = 3,$$
$$n_2 = 10, \qquad c_2 = 5.$$

Defina

X = Número de artículos defectuosos observados en la primera muestra.

Y = Número de artículos defectuosos observados en la segunda muestra.

Z = Número de artículos defectuosos observados en las muestras combinadas ($Z = X + Y$).

La curva OC es

$$\begin{aligned}\text{OC}(p) = p\{\text{Se acepta el lote} \mid p\} &= P\{\text{Se acepta el lote en la primera muestra} \mid p\} \\ &+ P\{\text{Se acepta el lote en la segunda muestra} \mid p\}\end{aligned}$$

donde

$$P\{\text{Se acepta el lote en la primera muestra} \mid p\} = P\{X \le 3 \mid p\}$$

y

$$\begin{aligned}&P\{\text{Se acepta el lote en la segunda muestra} \mid p\} \\ &= P\left\{\begin{matrix}\text{El lote ni se acepta ni se rechaza} \\ \text{en la primera muestra y el lote se} \\ \text{acepta en la segunda muestra}\end{matrix} \middle| \; p\right\} \\ &= P\{3 < X \le 5, Z \le 5 \mid p\}.\end{aligned}$$

El cálculo de esta probabilidad conjunta debe hacerse cuidadosamente, puesto que X y Z son variables aleatorias dependientes.

Considere $p = \text{AQL} = .1$.

$$\begin{aligned}&P\{\text{Se acepta el lote en la primera muestra} \mid p = .1\} \\ &= P\{X \le 3 \mid p = .1, n = 20\} = .8670,\end{aligned}$$

$$\begin{aligned}&P\{\text{Se acepta el lote en la segunda muestra} \mid p = .1\} \\ &= P\{X = 4 \mid p = .1, n = 20\}\, P\{Y \le 1 \mid p = .1, n = 10\} \\ &\quad + P\{X = 5 \mid p = .1, n = 20\}\, P\{Y \le 0 \mid p = .1, n = 10\} \\ &= (.0898)(.7361) + (.0319)(.3487) = .0772.\end{aligned}$$

Resumiendo:

$$P\{\text{Se acepta el lote} \mid p = .1\} = .8670 + .0772 = .9442.$$

La repetición de cálculos similares con $p = .3$ da

$$\begin{aligned}&P\{\text{Se acepta el lote} \mid p = .3\} \\ &= .1071 + (.1304)(.1493) + (.1789)(.0282) \\ &= .1316.\end{aligned}$$

Así, se concluye que para este caso obtenemos

$$\alpha = 1 - .9442 = .0558,$$
$$\beta = .1316.$$

La experimentación con otros valores de n_1, n_2, c_1, y c_2 puede conducir a planes de muestreo doble que concuerden aún más con los valores deseados de α y β. Se dispone de tablas para optimizar los planes de muestreo doble. (Vea, por ejemplo, Duncan, 1986, pp. 232-33.)

Problemas para la sección 11.10

34. Considere el plan de muestreo doble para Spire Records presentado en esta sección.

a) Suponga que la proporción verdadera de artículos defectuosos en el lote es 10 por ciento. En promedio, ¿cuántos artículos tendrán que muestrearse antes de que el lote sea aceptado o rechazado?

b) Suponga que la proporción verdadera de artículos defectuosos en el lote es 30 por ciento. En promedio, ¿cuántos artículos tendrán que muestrearse antes de que el lote sea aceptado o rechazado?

35. En el plan de muestreo doble para Spire Records presentado en esta sección, ¿cuál es la probabilidad de que se rechace un lote en la primera muestra? Realice el cálculo para $p = p_0$ y $p = p_1$.

36. Considere el plan de muestreo doble para Spire Records descrito en esta sección. Para un periodo de un año, 3 860 cajas de discos están sujetas a inspección usando este plan. Si el 60 por ciento de estos lotes son "buenos" (es decir, en el 60 por ciento de los lotes, la proporción de artículos defectuosos es exactamente 10 por ciento) y 40 por ciento son "malos" (es decir, en el 40 por ciento de los lotes, la proporción de artículos defectuosos es exactamente 30 por ciento), entonces ¿cuál es el número esperado de grupos

    a) aceptados?
    b) rechazados?
    c) aceptados en la primera muestra?
    d) aceptados en la segunda muestra?
    e) rechazados en la primera muestra?
    f) rechazados en la segunda muestra?

37. Grafique la curva OC para el plan de muestreo doble con $n_1 = 20$, $n_2 = 10$, $c_1 = 3$, y $c_2 = c_3 = 5$, como se describe en esta sección. Si usted está haciendo esto a mano, evalúe la curva para $p = 0$, .2, .4, .6, .8 y 1 solamente. (Sugerencia: la curva OC para este plan de muestreo tiene la forma

$$\begin{aligned}\text{OC}(p) = {} & P\{X \leq 3 \mid p, n = 20\} \\ & + P\{X = 4 \mid p, n = 20\}\, P\{Y \leq 1 \mid p, n = 10\} \\ & + P\{X = 5 \mid p, n = 20\}\, P\{Y = 0 \mid p, n = 10\}.)\end{aligned}$$

38. Por prueba y error diseñe un plan de muestreo doble para Spire Records que alcance $\alpha \approx .10$ y $\beta \approx .10$.

39. Considere el siguiente plan de muestreo doble. Primero seleccione una muestra de 5 de un lote de 100. Si hay cuatro o más artículos defectuosos en la muestra, rechace el lote. Si hay uno o menos artículos defectuosos, acepte el lote. Si hay dos o tres artículos defectuosos, muestre cinco artículos adicionales y rechace el lote si el número combinado de artículos defectuosos en ambas muestras es cinco o más. Si el lote tiene 10 artículos defectuosos, ¿cuál es la probabilidad de que el lote pase la inspección?

40. Para el plan de muestreo doble descrito en el problema 39, determine lo siguiente:

    a) La probabilidad de que se rechace el lote con base en la primera muestra.
    b) La probabilidad de que se rechace el lote con base en la segunda muestra.
    c) El número esperado de artículos muestreados antes de que el lote se acepte o se rechace.

11.11 PLANES DE MUESTREO SECUENCIAL

Los planes de muestreo doble pueden ampliarse a planes de muestreo triple, los cuales también pueden ampliarse a planes de orden superior. La conclusión lógica de este proceso es el plan de muestreo secuencial. En un plan secuencial, los artículos se muestran uno a la vez. Después de cada muestreo, se registran dos números: el número de artículos muestreados y el número acumulado de artículos defectuosos observados. Con base en estos números, se toma una de tres decisiones: (1) aceptar el lote, (2) rechazar el lote, o (3) continuar el muestreo. A diferencia de los planes de muestreo simple y doble, siempre existirá un plan de muestreo secuencial que dará valores específicos de p_0, p_1, α, y β. Los planes de muestreo secuencial se definen me-

diante tres regiones: la región de aceptación, la región de rechazo, y la región de muestreo. Las tres regiones están separadas por líneas rectas. Las líneas tienen la forma

$$L_1 = -h_1 + sn,$$
$$L_2 = h_2 + sn,$$

donde n es el número de artículos muestreados. Observe que L_1 y L_2 son ambas funciones lineales de la variable n. Las intersecciones de y son respectivamente $-h_1$ y h_2, y la pendiente de cada línea es s. Como las líneas tienen la misma pendiente, son paralelas. El plan de muestreo secuencial se implementa de la siguiente manera: se grafica el número acumulado de artículos defectuosos, junto con las líneas L_1 y L_2. Si el número acumulado de artículos defectuosos sobrepasa a L_2, se rechaza el lote, y si el número acumulado de artículos defectuosos se sitúa debajo de L_1, se acepta el lote. Siempre que el número acumulado de artículos defectuosos se sitúe entre L_1 y L_2, continúa el muestreo.

La figura 11-16 muestra dos ejemplos de los resultados del muestreo para el mismo plan de muestreo secuencial. En el caso A, el muestreo condujo a la aceptación del lote, y en el caso B condujo al rechazo del lote.

Las ecuaciones para h_1, h_2 y s son

$$h_1 = \frac{\log \dfrac{1-\alpha}{\beta}}{\log \dfrac{p_1(1-p_0)}{p_0(1-p_1)}},$$

$$h_2 = \frac{\log \dfrac{1-\beta}{\alpha}}{\log \dfrac{p_1(1-p_0)}{p_0(1-p_1)}},$$

$$s = \frac{\log \dfrac{1-p_0}{1-p_1}}{\log \dfrac{p_1(1-p_0)}{p_0(1-p_1)}}.$$

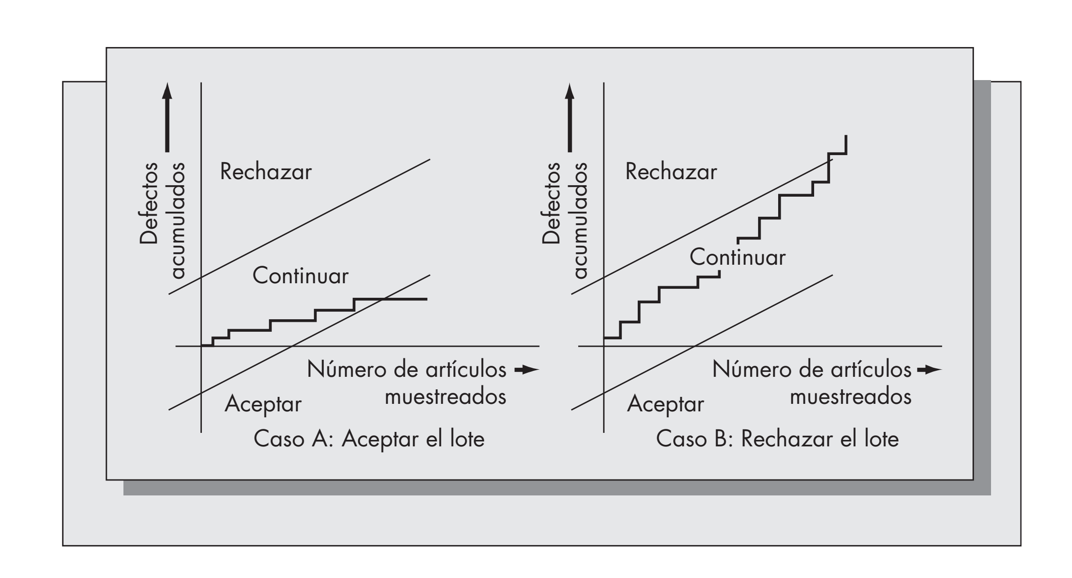


FIGURA 11-16
Dos ejecuciones de un plan de muestreo secuencial.

Ejemplo *11.8*

Considere nuevamente el ejemplo 11.6 de Spire Records. Herman Sondle ha experimentado con diferentes planes de muestreo para alcanzar los niveles deseados de los riesgos del consumidor y del productor. Decide construir un plan de muestreo secuencial para ver cómo se compara con los planes simple y doble que se presentaron anteriormente. Spire y B&G llegaron a un acuerdo con los siguientes valores de AQL, LTPD, riesgo del consumidor, y riesgo del productor:

$$p_0 = .1, \qquad \alpha = .1,$$
$$p_1 = .3, \qquad \beta = .1.$$

Observe que los denominadores en las expresiones para h_1, h_2, y s son los mismos. Primero calculamos el denominador. (Usaremos logaritmos de base 10 en nuestros cálculos. Ya que todas las fórmulas incluyen relaciones de logaritmos, los resultados serán los mismos independientemente de que se use base 10 o base *e*.)

$$\log[(.3)(.9)/(.1)(.7)] = 0.58626.$$

Así,

$$h_1 = \log(.9/.1)/0.58626 = 0.9542/0.58626 = 1.6277,$$
$$h_2 = h_1 = 1.6277 \qquad \text{(ya que } \alpha = \beta \text{ para este caso)},$$
$$s = \log(.9/.7)/0.58626 = 0.10914/0.58626 = 0.18617.$$

La figura 11-17 muestra las regiones de decisión para el plan de muestreo secuencial de Spire Records.

Cuando Herman sugirió que Spire Records implementara el plan de muestreo secuencial mostrado en la figura 11-17, encontró una resistencia considerable entre algunos de sus colegas de trabajo responsables de la inspección de las existencias que llegaban. Con un plan de muestreo simple de $n = 25$, argumentaron que cuando menos ellos sabrían con antelación cuántos discos tendrían que verificar. Sin embargo, con el plan secuencial, argumentaron que tal vez tendrían que muestrear el lote completo sin que el plan recomendara aceptación ni rechazo. Aunque Herman había oído que los planes secuenciales eran más eficientes, le fue difícil convencer a sus compañeros de trabajo para que hicieran la prueba con el plan.

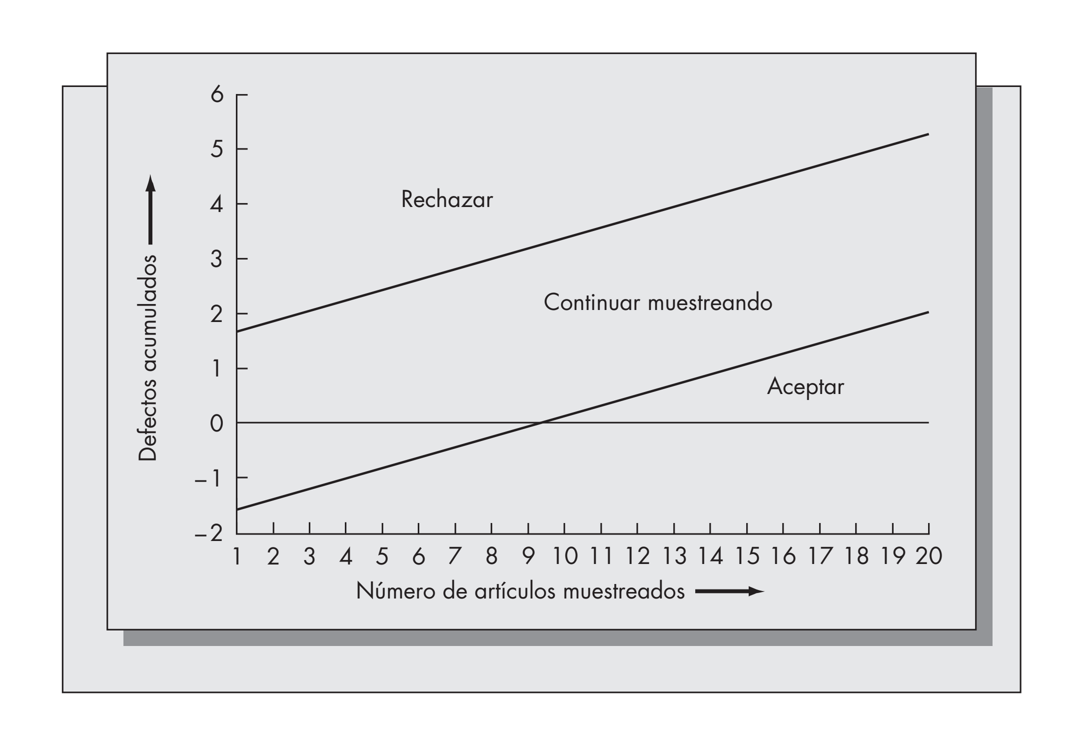


FIGURA 11-17
Plan de muestreo secuencial para Spire Records (consulte el ejemplo 11.8).

Los compañeros de trabajo en el ejemplo tenían razón al ver que el número de artículos muestreados cuando se usa el muestreo secuencial es una variable aleatoria. El tamaño de muestra *esperado* que resulta de un plan de muestreo secuencial depende de la proporción de artículos defectuosos en el lote, p. La curva del número promedio de muestras (ASN, *average sample number*) da el tamaño de muestra esperado para un plan de muestreo secuencial como función de p. Vamos a estimar la curva ASN obteniendo su valor para cinco puntos específicos: cuando $p = 0$, $p = p_0$, $p = s$, $p = p_1$, y $p = 1$. Es fácil encontrar la curva ASN para esos puntos. En la mayoría de los casos puede obtenerse una aproximación adecuada de la curva ASN cuando solamente su valor para estos cinco puntos. Los cinco valores son

$$\text{Para } p = 0, \qquad \text{ASN} = \frac{h_1}{s}.$$

$$\text{Para } p = p_0, \qquad \text{ASN} = \frac{(1-\alpha)h_1 - \alpha h_2}{s - p_0}.$$

$$\text{Para } p = s, \qquad \text{ASN} = \frac{h_1 h_2}{s(1-s)}.$$

$$\text{Para } p = p_1, \qquad \text{ASN} = \frac{(1-\beta)h_2 - \beta h_1}{p_1 - s}.$$

$$\text{Para } p = 1, \qquad \text{ASN} = \frac{h_2}{1-s}.$$

Considere la aplicación de estas fórmulas para Spire Records:

$$\text{ASN}(0) = \frac{1.6277}{0.18617} = 8.74,$$

$$\text{ASN}(.1) = \frac{(.9)(1.6277) - (.1)(1.6277)}{0.18617 - .1} = 15.11,$$

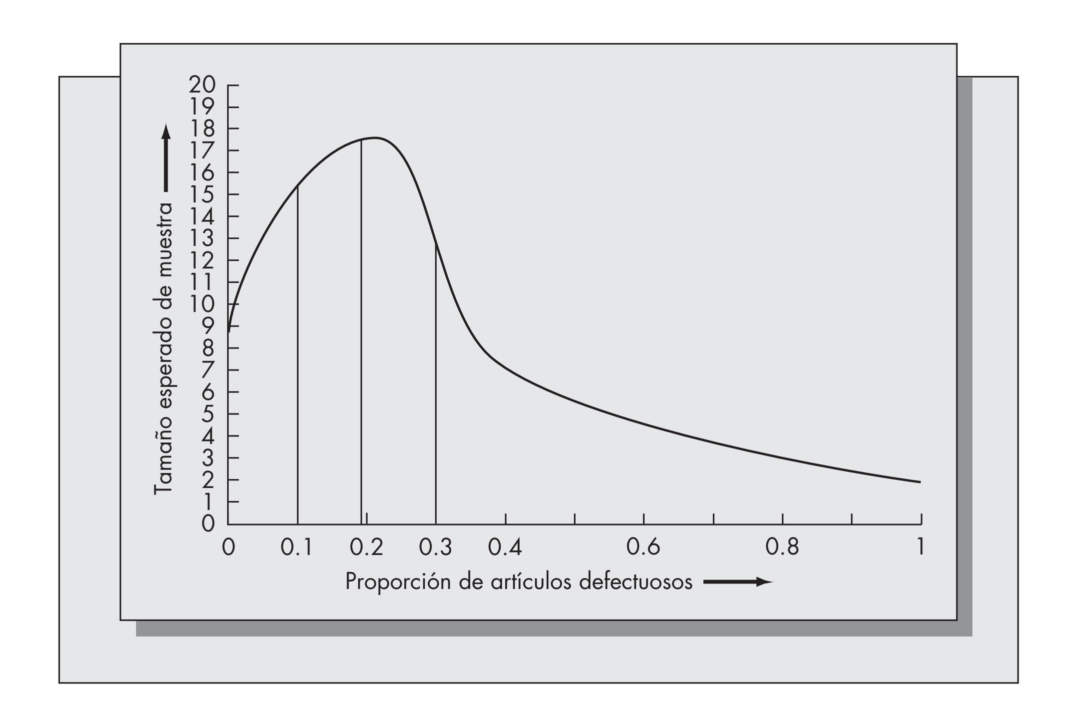
$$\text{ASN}(.18617) = \frac{(1.6277)(1.6277)}{(0.18617)(1 - 0.18617)} = 17.49,$$

$$\text{ASN}(.3) = \frac{(.9)(1.6277) - (.1)(1.6277)}{(.3 - 0.18617)} = 11.44,$$

$$\text{ASN}(1) = \frac{1.6277}{(1 - 0.18617)} = 2.0.$$

La curva ASN es una curva unimodal cuyo valor máximo se sitúa entre p_0 y p_1. Con base en esto y en los cinco puntos calculados, obtenemos la curva ASN estimada que se muestra en la figura 11-18. En la figura vemos que el tamaño esperado de muestra que el plan de muestreo secuencial para Spire Records requiere será cuando mucho de 18 artículos. Esto es evidentemente una mejora sobre el plan de muestreo simple con $n = 25$ y $c = 4$, que condujo a valores similares de α y β. Sin embargo, es importante tener en mente que el tamaño real de la muestra que se requiere para el plan secuencial con miras a inspeccionar cualquier grupo específico de discos es una variable aleatoria. La curva ASN da solamente el valor *esperado* de esta variable aleatoria para cualquier valor especificado de p. Así, es posible que en casos específicos, el tamaño real de la muestra pudiera ser mayor que 18, o aún mayor que 25.

FIGURA 11-18
Curva ASN para Spire Records (estimada).



Problemas para la sección 11.11

41. Un fabricante de motores de avión usa un plan de muestreo secuencial para aceptar o rechazar los lotes entrantes de microprocesadores que se usan en los motores. Suponga un AQL de 1 por ciento y un LTPD de 5 por ciento. Determine un plan de muestreo secuencial que suponga que $\alpha = .05$, $\beta = .10$. Grafique las regiones de aceptación y de rechazo.

42. Considere el plan de muestreo secuencial descrito en el problema 41. Suponga que se inspecciona un lote de 1 000 microprocesadores. Suponga que los chips probados 31, 89, 121 y 122 se encuentran defectuosos. Suponiendo el plan de muestreo secuencial obtenido en el problema 41, ¿será el lote aceptado, rechazado, o ninguno de los casos anteriores cuando se haya probado el chip 122?

43. Estime la curva ASN para el plan obtenido en el problema 41. Según su curva, ¿cuál es el número esperado de microprocesadores que deben probarse si la proporción verdadera de artículos defectuosos en el lote es

    a) 0.1 por ciento?

    b) 1.0 por ciento?

    c) 10 por ciento?

44. Considere el ejemplo de Hemispherical Conductor y de la compañía Sayle estudiado en el problema 32. Diseñe un plan de muestreo secuencial para Sayle que conduzca a $\alpha = .05$ y $\beta = .20$. ¿Cuáles son las ventajas y desventajas de este plan sobre el plan de muestreo Sayle obtenido en el problema 32?

45. Estime la curva ASN para el plan de muestreo obtenido en el problema 44. En promedio, ¿cuántos procesadores tendrían que probarse si

    a) $p = .001$?

    b) $p = .005$?

    c) $p = .01$?

11.12 CALIDAD PROMEDIO DE SALIDA

El propósito de un plan de muestreo es detectar los lotes de calidad inaceptable. Sin embargo, como el muestreo es un proceso estadístico, es posible que pasen los lotes malos y que se rechacen los lotes buenos. Un aspecto fundamental relacionado con la efectividad de cualquier plan de muestreo es determinar la calidad del producto que resulta *después* que se termina el proceso de inspección.

El cálculo de la calidad promedio de salida de un proceso de inspección depende del supuesto que se hace sobre los lotes que no pasan la inspección y del supuesto que se hace acerca de los artículos defectuosos. Suponga que los lotes rechazados están sujetos a una inspección al 100 por ciento. Obtenemos la curva de calidad promedio de salida bajo dos condiciones: (1) los artículos defectuosos en las muestras y en los lotes rechazados no se reemplazan y (2) los artículos defectuosos en las muestras y en los lotes rechazados se reemplazan.

La calidad promedio de salida (AOQ) es el cociente a largo plazo del número esperado de artículos defectuosos entre el número esperado de artículos que pasaron la inspección con éxito. Es decir,

$$\text{AOQ} = \frac{E\{\text{número de artículos defectuosos a la salida}\}}{E\{\text{número de artículos a la salida}\}}.$$

La curva OC es la probabilidad de aceptación de un lote como función de p. Es decir,

$$\text{OC}(p) = P\{\text{el lote se acepta} \mid p\}.$$

Por comodidad vamos a llamar a este término P_a.

Caso 1: Los artículos defectuosos no se remplazan. Suponga que los lotes tienen tamaño N y que las muestras son de tamaño n. Entonces el número esperado de artículos defectuosos y el número esperado de artículos fletados son

		Número de artículos defectuosos	Número de artículos
Aceptar el lote	P_a	$(N - n)p$	$N - np$
Rechazar el lote	$1 - P_a$	0	$N(1 - p)$

A partir de este diagrama de árbol vemos que

$$\begin{aligned} E\{\text{número de artículos defectuosos de salida}\} &= P_a(N - n)p + (1 - P_a)(0) \\ &= P_a(N - n)p \end{aligned}$$

y

$$E\{\text{número de artículos de salida}\} = P_a(N - np) + (1 - P_a)N(1 - p).$$

Se concluye que el cociente AOQ está dado por

$$\text{AOQ} = \frac{P_a(N - n)p}{P_a(N - np) + (1 - P_a)N(1 - p)} = \frac{P_a(N - n)p}{N - np - p(1 - P_a)(N - n)}.$$

Si $N \gg n$ (N es mucho mayor que n), que es un supuesto común, esta ecuación es aproximadamente

$$\text{AOQ} \approx \frac{P_a p}{P_a + (1 - P_a)(1 - p)} = \frac{P_a p}{1 - p(1 - P_a)}.$$

Las fórmulas son un poco más sencillas si los artículos defectuosos se reemplazan con artículos buenos.

Caso 2: *Los artículos defectuosos se reemplazan.* En este caso el diagrama de árbol se convierte en

		Número de artículos defectuosos	Número de artículos
Aceptar el lote	P_a	$(N - n)p$	N
Rechazar el lote	$1 - P_a$	0	N

La AOQ está dada por

$$\text{AOQ} = \frac{P_a(N - n)p}{N},$$

que es aproximadamente

$$\text{AOQ} \approx P_a p = \text{OC}(p)p$$

cuando $N \gg n$.

Esta última fórmula es la que con mayor frecuencia se usa en la práctica, principalmente por su simplicidad. Una medida importante de qué tan efectivo resulta un plan de muestreo es el *límite de calidad promedio de salida* (AOQL, *average outgoing quality limit*), que se define como el valor máximo de la curva AOQ.

Ejemplo *11.9* Considere el caso de Spire Records. En el caso de Spire podemos suponer que los lotes son grandes en comparación con las muestras y que todos los artículos defectuosos se remplazan. En ese caso AOQ ≈ OC(p)p. En la figura 11-19 hemos generado las curvas AOQ para los planes de muestreo simple de Spire con n = 10 y n = 25. Observe cómo el tamaño mayor de la muestra mejora significativamente la calidad promedio de salida. De la figura 11-19 vemos que si Spire usa un plan de muestreo simple con n = 25 y n = 4, la tienda puede esperar que la proporción de discos defectuosos B&G en sus anaqueles no sea mayor que aproximadamente el 10.2 por ciento.

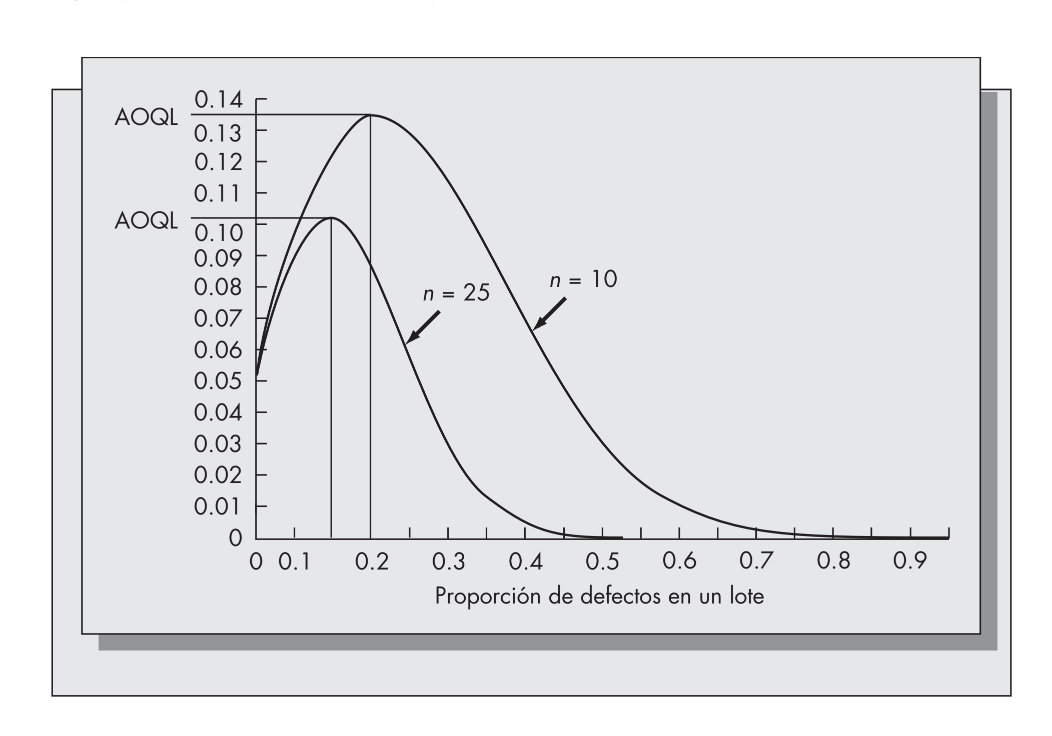


FIGURA 11-19
Curvas AOQ para Spire Records (consulte el ejemplo 11.9).

Aplicación instantánea

MOTOROLA LLEVA LA DELANTERA CON LOS PROGRAMAS DE CALIDAD SEIS-SIGMA

La Corporación Motorola ha logrado un impresionante registro en la definición e implementación de nuevas iniciativas de calidad y en la traducción de esas iniciativas en ganancias. Como vemos en este capítulo, los métodos tradicionales de calidad suponen que una condición fuera de control corresponde a una observación que se sitúa fuera de $\pm 2\sigma$ o $\pm 3\sigma$. Motorola decidió que dicho estándar era demasiado laxo, ya que daba demasiados defectos. En la década de los 80, se movieron hacia un estándar de 6σ. Para alcanzar esta meta, Motorola estableció la práctica de las revisiones del sistema de calidad (QSR por sus siglas en inglés) y puso mucho énfasis en las técnicas clásicas del control estadístico de procesos (SPC por sus siglas en inglés).

Para alcanzar el estándar 6σ, Motorola inculcó la filosofía de la gestión de la calidad total en toda su organización. Los programas de calidad no se asignaron a un solo grupo ni tampoco se convirtieron en su responsabilidad, sino que se convirtieron en parte del trabajo de cada uno. El enfoque de Motorola se basó en los siguientes ingredientes claves:

- El objetivo principal es la satisfacción total del cliente.
- Medición uniforme de la calidad para todas las partes del negocio.
- Expectativas consistentes de mejora en toda la compañía.
- Planes de incentivos orientados hacia objetivos para la gerencia y para los empleados.
- Programas coordinados de entrenamiento.

Motorola ha documentado cuidadosamente el mapa de caminos que conduce a sus metas de calidad (Motorola, 1993). El primer paso del proceso es una auditoría detallada de varias partes clave del negocio. Se incluye el control del desarrollo de los nuevos productos, el control de los proveedores (tanto internos como externos), el monitoreo de los procesos y del equipo, las consideraciones de recursos humanos, y la evaluación de la satisfacción del cliente. Aun la gerencia general toma parte en muchas de estas auditorías mediante visitas regulares a los clientes, dirigiendo juntas de los comités de la política de operación, y reconociendo a los ejecutivos que han hecho contribuciones extraordinarias a las iniciativas de calidad de la compañía.

Kumar y Gupta (1993) reportan sobre su experiencia con un programa TQM puesto en marcha en la planta de ensamblado de Motorola en Austin, Texas. En mayo de 1988, la gerencia comenzó el proceso de implementación de un programa CEP en Austin. El proceso comenzó al llamar a un consultor externo para que diseñara el programa, y asignando un coordinador interno para que finalmente tomara las responsabilidades del consultor. Para asegurar que los empleados adoptaban la iniciativa como propia, la gerencia organizaba equipos participativos para solucionar problemas. Cada equipo incluía un gerente de manufactura, un líder de grupo, operadores de los dos turnos, un representante del Departamento de Aseguramiento de Calidad, y un ingeniero. Austin tenía seis equipos en total. Para asegurar aun más que todos los empleados se comprometieran, le gerencia inició un programa de entrenamiento en SPC en toda la planta. El entrenamiento se adaptó a la función del puesto de trabajo.

El Departamento de Aseguramiento de Calidad de la planta instituyó un programa de certificación en Austin para los vendedores. Como resultado, aproximadamente el 60 por ciento de los vendedores que proveían a la planta fueron certificados. Dentro de la planta, se emplearon métodos tradicionales de SPC: se recabaron y graficaron datos de los atributos y las máquinas se apagaban cuando se detectaban situaciones fuera de control. Los miembros del equipo de aseguramiento de calidad emplearon el diseño de las técnicas de experimentación para identificar las causas de los problemas.

¿Cuál es el resultado? Durante los dos primeros años de esta iniciativa, la planta de Austin reportó una disminución del 56 por ciento de las tasas de merma. Al desarrollar una estrategia coordinada y claramente definida, Motorola tuvo la capacidad de alcanzar mejoras importantes en las mediciones tradicionales de la calidad en esta instalación. El éxito generalizado de Motorola es un testimonio de que éste no fue un ejemplo aislado. Demuestra que las compañías estadounidenses pueden competir de una manera efectiva con los competidores del extranjero si el esfuerzo de calidad es una iniciativa real en toda la compañía.

Problemas para la sección 11.12

46. Si se remplazan los artículos defectuosos y $N \gg n$, demuestre mediante el cálculo diferencial que el valor de p para el cual AOQ(p) alcanza su valor máximo satisface

$$\frac{d\mathrm{OC}(p)}{} = -\frac{\mathrm{OC}(p)}{}.$$

47. Considere el plan de muestreo simple con $n = 10$ y $c = 0$.

    a) Obtenga una ecuación analítica para la curva OC como función de p.
    b) Usando los resultados del problema 46, determine el valor de p para el cual la curva AOQ tiene un máximo.
    c) Usando los resultados de los incisos *a* y *b*, determine el valor máximo de la calidad promedio de salida.

48. Considere el plan de muestreo simple estudiado en el problema 30. Si se reemplazan los artículos defectuosos y $N \gg n$, grafique la curva AOQ y determine el valor de AOQL.

11.13 ADMINISTRACIÓN DE LA CALIDAD TOTAL

Este capítulo revisó los fundamentos del control estadístico de la calidad. El control estadístico de la calidad constituye un conjunto de técnicas basadas en las teorías de la probabilidad y en el muestreo estadístico para controlar la variación de los procesos y para determinar si los lotes de manufacturas cumplen con los niveles deseados de calidad. Sin embargo, la entrega de calidad al cliente es un problema mucho más amplio que los que se encaran con aspectos estadísticos solamente. Esta sección considera la calidad desde la perspectiva administrativa.

Definiciones

¿Qué es la administración de la calidad total (TQM, por el inglés *Total Quality Management*)? El término parece haber sido acuñado por primera vez gracias a Feigenbaum (1983; en una edición anterior), quien dio la siguiente definición:

> El control de la calidad total es un sistema efectivo para integrar los esfuerzos del desarrollo, el mantenimiento, y la mejora de la calidad de los diferentes grupos en una organización, lo que permite satisfacer completamente al cliente en términos de mercadotecnia, ingeniería, producción, y servicio.

El enfoque de Feigenbaum es definir la calidad en términos del cliente. Como observamos en la introducción de este capítulo, la mayoría de las definiciones de la calidad se refieren ya sea al apego a las especificaciones o a la satisfacción del cliente. Garvin (1988) amplía estas ideas y sugiere considerar la calidad a lo largo de ocho dimensiones básicas:

- Desempeño
- Características
- Confiabilidad
- Apego
- Durabilidad
- Facilidad de servicio
- Estética
- Calidad percibida

Podríamos agrupar las cinco primeras dimensiones bajo el encabezado general de apego a los requerimientos (la definición sugerida por Crosby, 1979), y las tres últimas bajo el encabezado de la satisfacción del cliente (como sugiere Feigenbaum, 1983). Sin embargo, al dividir estas dos categorías, Garvin aprecia mejor la complejidad de la calidad.

Escuchar al cliente

Un aspecto importante al diseñar productos de calidad es dar a las personas lo que quieren. Una cafetera perfectamente diseñada y construida que se venda en un lugar donde nadie bebe café

es un fracaso por definición. Así, parte del proceso de entregar calidad al cliente es saber lo que quiere.

Aunque escuchar al cliente es una parte importante para enlazar el ciclo manufactura/diseño, generalmente se le asocia mucho más con la mercadotecnia que con las operaciones. Aun así, según vemos, los límites que separan las áreas funcionales de las de negocios se hacen cada vez más difusos. La manufactura no puede operar en el vacío. Debe ser parte del enlace con el cliente.

Encontrar lo que quiere el cliente e incorporar esos deseos en el diseño y la manufactura del producto es un proceso de pasos múltiples. Los pasos de tal proceso son

- Obtención de los datos.
- Caracterización de las necesidades del cliente.
- Priorización de las necesidades del cliente.
- Enlace de las necesidades con el diseño.

Hay varios medios para obtener los datos iniciales. Tradicionalmente, la opinión del cliente se solicita a través de entrevistas y encuestas. Hay muchos aspectos que deben tomarse en cuenta al considerar las entrevistas con los clientes o con los clientes potenciales. ¿Cuantas respuestas de clientes son suficientes? La respuesta correcta varía según factores distintos. ¿Cuántos segmentos de mercado existen para el producto? ¿Cuántos atributos son importantes? ¿Qué métodos van a usarse para interpretar los resultados? Enseguida, surge la interrogante de cómo solicitar la información al cliente. ¿Deben realizarse entrevistas o encuestas? La respuesta no es clara. Ambas tienen ventajas. Las entrevistas permiten respuestas más abiertas, pero los prejuicios del entrevistador podrían sesgar los resultados. La redacción de las preguntas determina tanto las encuestas como las entrevistas. Por ejemplo, suponga que Mr. Coffee está considerando un nuevo diseño para una cafetera. Una pregunta como "¿Cuál debe ser la capacidad de una cafetera automática?" supone de inmediato que el cliente se preocupa por la capacidad. La pregunta "¿Prefiere una cafetera de 8 o de 12 tazas?" impone aún más supuestos (Dahan, 1995).

Los grupos de enfoque son otra popular técnica para solicitar la voz del consumidor. El formato del grupo de enfoque tiene la ventaja de ser abierto; la redacción específica de las preguntas no es tan importante como con las encuestas o las entrevistas. Sin embargo, los grupos de enfoque tienen desventajas. El moderador puede afectar el flujo de la discusión. También es probable que los participantes con personalidad fuerte dominen al grupo.

Una vez que se desarrolla la base de datos, deben priorizarse y agruparse las necesidades y los deseos del cliente. Se dispone de varios métodos para esto. Uno que ha recibido mucha atención en la literatura de mercadotecnia es el análisis conjunto (debido a Green y Rao, 1971). El análisis conjunto es una técnica basada en la estadística para estimar la utilidad de los atributos del producto con base en los datos de la preferencia del cliente.

Una vez que se determinan y se agrupan los atributos, es necesario enlazar estos atributos con los procesos de diseño y manufactura. Esto puede hacerse desplegando la función de calidad (QFD, por sus siglas en inglés). Con QFD, las necesidades del cliente se relacionan con los atributos del producto y/o los aspectos del proceso de producción mediante una matriz. El usuario proporciona estimaciones de la correlación entre los atributos y las necesidades en la parte del "techo" de la matriz. La figura que resulta se parece un poco a una casa, de ahí el término *casa de la calidad* para describir la matriz resultante (véase la figura 11-20). QFD se usa conjuntamente con los métodos tradicionales tales como las encuestas y los grupos de enfoque. La fuerza de las correlaciones entre las necesidades del cliente y los atributos del producto o las características de diseño muestran dónde debe ponerse el énfasis cuando se considera el diseño de productos nuevos o los cambios de diseño en los productos existentes. El lector que tenga interés debe consultar a Cohen (1995) para un estudio exhaustivo actualizado de los métodos QFD.

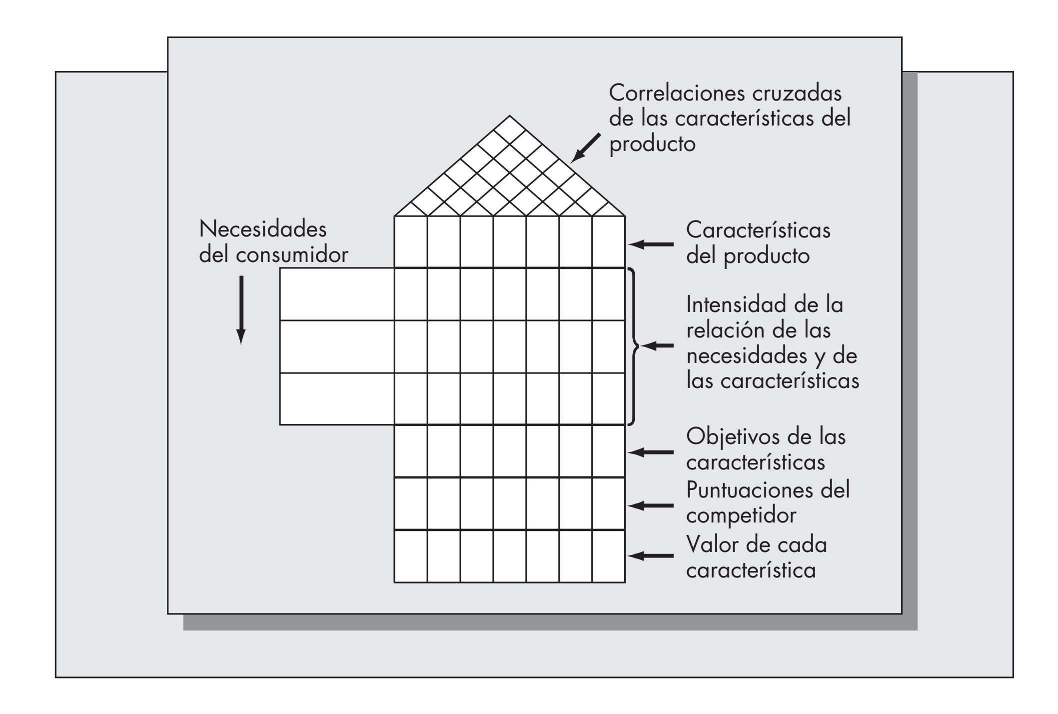


FIGURA 11-20
La casa de la calidad: la matriz de planeación QFD.

Esta sección solamente tocó la superficie de la solicitud de la opinión del cliente y de su integración en el proceso del diseño y la manufactura del producto. Sin embargo, debe tenerse cuidado en no poner *demasiado* énfasis en la voz del cliente. Las innovaciones reales no provienen del cliente sino de los visionarios. Al dejar que el consumidor sea el rey, la innovación podría sofocarse. De acuerdo con Martin (1995):

> Todavía, más y más compañías están aprendiendo que algunas veces sus clientes en realidad pueden extraviarlo. El peligro radica en convertirse en un fanático de la retroalimentación, servilmente dedicado a los clientes, tratando constantemente de tener mejores contactos a través de más grupos de enfoque, más encuestas. Eso no solamente puede distraerlo del trabajo real a la mano, sino que también puede moverlo a crear nuevas ofertas que son seguras y blandas.

Productos nuevos de éxito memorable como las notas Post-it^{MR} desarrolladas por 3M no fueron el resultado de las encuestas de mercadotecnia. Muchos productos que parecían ganadores en la investigación de mercado se desplomaron en el mercado real. Un ejemplo es la Nueva Coca, que ganó indiscutiblemente en las pruebas de sabor pero nunca fue aceptada por los consumidores. Por supuesto que probablemente hay más ejemplos sobre desplomes que resultaron de no escuchar al cliente. La moraleja aquí es que aun cuando el cliente es una parte importante de la ecuación, las encuestas de la clientela y similares no deben usarse para excluir el ingenio y la visión.

Competencia basada en la calidad

Ya que es posible definir la calidad de varias maneras, frecuentemente la gerencia debe escoger a lo largo de cuáles dimensiones de calidad va a enfocarse. La Corporación Cray desarrolló una reputación como productor de las computadoras más rápidas disponibles. Para lograrlo, Cray depositó sus energías en la tecnología más que en la confiabilidad. Citando al gerente de la compañía: “Si una máquina no se descompone una vez al mes o por el estilo, es probable que no esté optimizada apropiadamente” (como lo relata T. Alexander, 1985). Según observa Garvin (1988), este enfoque permitió a Cray ser un líder tecnológico, pero lo hizo vulnerable a una competencia con un desempeño comparable y una mejor confiabilidad.

Otras compañías han escogido diferentes dimensiones de calidad mediante las cuales competir. El margen competitivo de las computadoras Tandem se basa en la confiabilidad del producto. Con los procesadores paralelos, Tandem está en condiciones de garantizar que esencialmente no hay tiempo de inactividad. El enfoque de Tandem ha tenido gran éxito porque muchos clientes, como los bancos y las dependencias de gobierno, están dispuestos a pagar una bonificación por mejorar la confiabilidad.

La alta confiabilidad también es uno de los principales factores del éxito de los fabricantes de automóviles japoneses. Con un mayor número de familias con carreras duales, la confiabilidad es muy difícil de conseguir. Con base en la encuesta de los lectores anuales de la Unión de Consumidores, los fabricantes japoneses de automóviles superan de manera consistente a sus competidores estadounidenses y europeos en esta dimensión de la calidad.

La banca es una industria que compite a lo largo de la dimensión del servicio. Por ejemplo, el banco Wells Fargo de California proporciona a sus clientes una gran red de sucursales, muchos cajeros automáticos, y otros servicios como información actualizada sobre su cuenta 24 horas diarias. Los bancos más pequeños atraen a los clientes mediante tasas de interés más altas en los ahorros, tasas más bajas en los intereses de las tarjetas de crédito, y servicio personalizado.

Algunas compañías han basado sus estrategias en suministrar al consumidor el producto de la más alta calidad, independientemente del costo. Ejemplos típicos son los relojes Rolex, las cámaras Leica, los automóviles Rolls Royce, las plumas Cross, y los pianos Steinway. Muchas compañías estadounidenses prefieren dirigir sus productos al mercado masivo y dejar el mercado más pequeño de alto nivel a los competidores europeos y japoneses.

Varias compañías exitosas dependen de una estrategia de negocios basada en llevar el liderazgo del mercado a lo largo de una o dos dimensiones de calidad. Sin embargo, esto puede hacerlos vulnerables a la competencia.

Organización para la calidad

TQM requiere una dedicación total por parte de la gerencia y los trabajadores. La organización necesita estructurarse para magnificar la calidad del producto, no desmerecerla. Los trabajadores deben tener una posición segura y participar del éxito de la organización si los programas de calidad van a tener éxito. Según Feigenbaum (1983):

1. La función de calidad dentro de una compañía no debe ser una función individual alojada en un solo departamento. La calidad debe reconocerse como un grupo sistemático de disciplinas que se aplicarán sobre una base coordinada por todas las funciones a lo largo de la compañía y de la planta.
2. La función de calidad debe tener un contacto directo y continuo con los compradores y los clientes de los productos y los servicios de la compañía.
3. La función de calidad debe organizarse para trascender los límites organizacionales de la función individual.
4. La función general de calidad debe supervisarse desde un alto nivel de la compañía a medida que se desarrollan nuevos productos, para asegurar que los aspectos de calidad se tratan adecuadamente, que se reconocen y se corrigen las "alarmas tempranas" de los problemas inminentes, y que la gerencia puede administrar apropiadamente la calidad.

¿Cuáles son las causas básicas de los problemas de calidad y qué puede hacerse para encarar estas causas? Leonard y Sasser (1982) encuestaron a ejecutivos de 30 compañías de *Fortune 500* para determinar sus opiniones en cuanto a las causas de los problemas de calidad en Estados Unidos. El factor en la parte superior de su lista era la mano de obra y la destreza del trabajador. Probablemente sea justo suponer que una encuesta similar de los trabajadores apuntaría hacia la gerencia. Las políticas de la gerencia determinan la seguridad en el empleo de los trabajadores y las condiciones de trabajo. Las estructuras de recompensa determinan los incentivos por el comportamiento. Para que TQM funcione, la organización completa debe alinearse detrás de un imperativo de la calidad. Para que esto ocurra, la gerencia debe ser responsable

de crear una estructura organizativa en la cual los trabajadores tengan poderes y en la cual los trabajadores tengan incentivos para hacer su mejor esfuerzo, no una que asfixie la excelencia.

Un programa que recibió mucha atención en los inicios de la década de los 80s eran los *círculos de calidad* (QC). Los círculos de calidad fueron una tentativa por emular la estructura organizativa japonesa. Un círculo de calidad es un pequeño grupo de aproximadamente 6 a 12 empleados que se reúnen regularmente con el propósito de identificar, analizar, y sugerir soluciones para los problemas que afectan su área de trabajo. Comúnmente, un grupo de este tipo se reuniría durante aproximadamente cuatro horas al mes en horas de trabajo. Al grupo puede dársele entrenamiento especial sobre los métodos del control estadístico de la calidad, dinámicas de grupo, y técnicas generales de solución de problemas.

Hubo un incremento dramático en el interés sobre los círculos de calidad en Estados Unidos que comenzó aproximadamente en 1980. Lawler y Morhman (1985) reportan que 44 por ciento de las compañías en Estados Unidos con más de 500 empleados tuvieron algún tipo de programa de círculo de calidad puesto en 1982 y que el 75 por ciento de estos programas inició después de 1980. Hay varias razones para el interés repentino en los círculos de calidad en Estados Unidos:

1. *Costo*. Los círculos de calidad pudieron instalarse de manera relativamente barata. Las compañías consultoras que se especializan en el área instalaban el programa, entrenaban a las personas apropiadas, y supervisaban las actividades iniciales del círculo. Desde la perspectiva de los costos, una ventaja del círculo de la calidad es que no necesariamente intervienen todos los empleados de la compañía.
2. *Control*. Desde el punto de vista del gerente, una característica deseable del círculo de calidad es que no tiene un poder formal para tomar decisiones, y por lo tanto no se percibe como una amenaza. El círculo de calidad sirve esencialmente para jugar el papel de un mecanismo formal de "buzón de sugerencias", y no hay ningún requerimiento de que el gerente renuncie a ninguna de las prerrogativas de la gerencia en su área de responsabilidad.
3. *Moda.* Hay poca duda de que los círculos de calidad han tenido una época de enorme atractivo. El éxito evidente de los programas de este tipo en la industria japonesa ha conducido a una aceptación anticipada de los QC en Estados Unidos.

Con el interés en los círculos de calidad y la adopción de tantos programas de QC en Estados Unidos, ¿qué impacto han tenido estos programas? La experiencia en todos sentidos ha sido que pocos programas de QC han evolucionado hacia otros programas que podrían afectar los procedimientos y las practicas de la compañía. Como resultado, el interés de grupo en los QC ha tendido a disminuir y los grupos se han reunido con menos frecuencia o ya no se han reunido. Lawler y Mohrman (1985) sugieren que el QC es inherentemente una estructura organizativa inestable. Como no tiene poder para tomar decisiones, su único papel es servir como un "buzón de sugerencias" formal. Es natural que los empleados finalmente perdieran el interés. En última instancia, los círculos de calidad pudieron haber servido como un primer paso hacia la construcción de un enfoque donde la gerencia participara más, pero eso no parece haber ocurrido.

Benchmarking para la calidad

El *benchmarking* implica medir el rendimiento propio contra el de los competidores. El benchmarking competitivo está ganando importancia a la luz del incremento de la competencia global. Resulta esencial establecer intereses particulares y cerciorarse de que estas prioridades son consistentes con las necesidades del mercado. Tomando como punto de partida una base de datos desarrollada en la Universidad de Boston, Miller *et al.* (1992) reportan prioridades competitivas en Europa, Japón y Estados Unidos. Éstas se resumen en la tabla 11-6.

Hay varias cosas interesantes que pueden verse en esta tabla. Primero, consistente con nuestro análisis, las compañías japonesas encuestadas le asignan a la confiabilidad del producto la prioridad máxima. Los japoneses entienden que la confiabilidad del producto podría ser

TABLA 11-6 Las cinco primeras prioridades competitivas

Europa	Japón	Estados Unidos
Calidad de apego	Confiabilidad del producto	Calidad de apego
Entrega a tiempo	Entrega a tiempo	Entrega a tiempo
Confiabilidad del producto	Cambio rápido de diseño	Confiabilidad del producto
Calidad del desempeño	Calidad de apego	Calidad del desempeño
Velocidad de entrega	Producto hecho a la medida	Precio

Fuente: Miller *et al*., 1992.

su mayor carta competitiva y planean seguir enfatizando esta importante dimensión de la calidad. También es interesante ver que solamente las compañías estadounidenses mencionan el precio.

Los autores listan cuatro tipos de benchmarking:

1. Benchmarking del producto.
2. Benchmarking funcional o del proceso.
3. Benchmarking de las mejores prácticas.
4. Benchmarking estratégico.

El *benchmarking del producto* significa desarmar el producto de un competidor para ver qué puede aprenderse sobre su diseño y construcción. Se dice que cuando Toyota inició su programa para producir el Lexus previendo competir con automóviles tales como Mercedes y BMW, examinó cuidadosamente los productos del competidor para determinar cómo y dónde se colocaban las soldaduras, y cómo se armaban los automóviles para alcanzar la apariencia y el tacto de calidad excepcional.

El *benchmarking funcional* se enfoca en el proceso más que en el producto. Los procesos típicos podrían ser la entrada de la orden, el ensamblado, el desarrollo de pruebas del producto, y el flete. El benchmarking funcional es posible sólo cuando las compañías tienen la voluntad de cooperar y compartir la información. Tiene el mismo objetivo que el benchmarking del producto: mejorar el proceso y finalmente el producto resultante.

El *benchmarking de las mejores prácticas* es similar al *benchmarking* funcional, salvo que se enfoca en las prácticas gerenciales en vez de un proceso específico. Las mejores prácticas pueden considerar factores como el ambiente de trabajo y los incentivos salariales para los empleados en compañías con un desempeño excepcional. General Electric es un fuerte partidario del benchmarking de las mejores prácticas (*Fortune*, 1991).

El objetivo de la *referencia estratégica* es considerar los resultados de otras comparaciones de benchmarking a la luz del enfoque estratégico de la compañía. Específicamente, ¿cuál es la estrategia general de negocios que ha sido articulada por el CEO, y son consistentes con esta estrategia los resultados de otros estudios de benchmarking?

En última instancia, ¿cual es el propósito del benchmarking? Es asegurar la mejora continua y es sólo uno de los medios para alcanzarla. La mejora continua del producto y del proceso es la meta última de cualquier programa de calidad. El benchmarking competitivo proporciona una manera de aprender de nuestros competidores. Aunque el benchmarking puede ser una herramienta útil, no es un sustituto de una estrategia de negocios claramente articulada ni de una visión para la compañía.

El Premio Deming y el Premio Baldrige

Los tres líderes del movimiento de calidad en Estados Unidos durante la década de los 50, W. Edwards Deming, Joseph M. Juran, y A. V. Feigenbaum, cada uno contribuyó al movimiento japonés de la calidad, aunque Deming es ciertamente el nombre que primero viene a la mente.

El éxito de Deming en Japón fue resultado de un seminario que presentó sobre control estadístico de la calidad en 1950. Repitió ese seminario varias veces, se hizo activo en el movimiento japonés de la calidad y finalmente se convirtió en un héroe nacional en Japón. Deming recomendaba aplicar tanto los métodos estadísticos como un enfoque sistemático para resolver los problemas de calidad. Posteriormente su enfoque llegó a conocerse como el método Planear, Hacer, Verificar, Actuar (PDCA).

Juran, otro líder importante del movimiento de la calidad, enfatizó los aspectos gerenciales de la calidad más que los aspectos estadísticos. Juran también presentó varios seminarios a los japoneses que estaban dirigidos desde a la gerencia media hasta la superior. El Instituto Juran, ubicado en Wilton, Connecticut, se fundó en 1979 y sigue proveyendo consultoría y entrenamiento en el área de la calidad. [Sin embargo, Philip Crosby y asociados, fundado el mismo año, genera casi 10 veces el ingreso anual (*Business Week*, 1991).]

Usando las regalías de un libro basado en sus conferencias de 1950, Deming estableció un premio nacional de la calidad en Japón. Los fondos para el premio siguen proviniendo principalmente de las regalías de Deming, pero están suplementados por una compañía periodística de Japón y por donaciones privadas. El premio Deming se otorga a (1) aquellos que han alcanzado la excelencia en la investigación en la teoría o en la aplicación del control estadístico de la calidad, (2) aquellos que han hecho contribuciones notables a la difusión de los métodos del control estadístico de la calidad, y (3) a aquellas corporaciones que han logrado resultados encomiables en la práctica del control estadístico de la calidad.

El premio ha sido otorgado casi cada año en varias categorías desde 1951. Los galardonados incluyen compañías bien conocidas como Toyota Motor Co., Hitachi, Fuji Film, Nippon Electric y NEC. También se han otorgado premios a divisiones específicas de compañías grandes y a plantas específicas dentro de una división. Es un honor nacional altamente codiciado en Japón (Aguayo, 1990).

El Reconocimiento a la Calidad Nacional Malcolm Baldrige fue establecido por el Departamento de Comercio de Estados Unidos en 1987 mayormente como una respuesta al éxito del premio Deming en Japón. El reconocimiento lleva el nombre del extinto secretario de Comercio, quien murió en un accidente el mismo año que el Congreso y el presidente Ronald Reagan promulgaron la Ley de la Mejora de la Calidad Nacional Malcolm Baldrige.

El reconocimiento se otorga cada año en tres categorías: (1) las compañías manufactureras o subsidiarias, (2) las compañías de servicio o subsidiarias, y (3) los pequeños negocios. La solicitud del reconocimiento Baldrige es un proceso complicado que requiere una dedicación sincera por parte de la compañía que busca el reconocimiento. Un comité de supervisores designados por el secretario de comercio tiene la voz final. Ellos basan su evaluación en el siguiente sistema de valores de nueve puntos:

1. Administración de la calidad total.
2. Utilización de los recursos humanos.
3. Desempeño.
4. Aspectos medibles.
5. Satisfacción del cliente.
6. Calidad de clase mundial.
7. Calidad temprana en el proceso.
8. Innovación.
9. Liderazgo externo.

El proceso de evaluación utiliza un complicado sistema de puntuación e incluye visitas al sitio por parte del comité de evaluación. Los fondos para apoyar el proceso de examen son donados por personas y compañías. Para el 30 de junio de 1989, se habían reunido más de $10.4 millones para apoyar al programa (Pierce, 1991). El número de solicitantes creció cada año

durante los primeros tres años después que el premio se anunció. En 1989, 40 compañías solicitaron los reconocimientos; en 1990, fueron 97 compañías; en 1991, 106; y en 1992, 90. [Hay varias razones de la caída de las solicitudes en 1992. Una es que muchas compañías están más interesadas en la certificación ISO 9000. Otra es que el costo de solicitud es un problema para algunas compañías. Estos costos no son insignificantes. Por ejemplo, Xerox, el ganador de 1989, gastó $1 millón en preparar su solicitud. Finalmente, al verse forzados a despedir trabajadores, muchas compañías no están en condiciones como para desarrollar una aplicación creíble (Hillkirk, 1992).]

¿Cuál ha sido la experiencia con el reconocimiento Baldrige ahora que han transcurrido más de 10 años desde su inicio? Una manera de medir el éxito del programa es ver cómo se han comportado las compañías que han ganado. Aun cuando no todos los ganadores Baldrige han sido ganadores en el mercado, la evidencia es que en promedio, las compañías parecen superar a la economía como un todo. De acuerdo con un experimento realizado por el Instituto Nacional de Estándares y Tecnología (NIST), cinco compañías ganadoras —Eastman Chemical, Federal Express, Motorola, Solectron, y Zytec— han superado el Standard & Poors 500 por 6.5 a 1. Las compañías matrices comercializadas públicamente de siete filiales que ganaron el premio superaron el S&P 500 por casi 3 a 1 (*Transportation and Distribution*, 1995). Aun cuando no todos los ganadores Baldrige hicieron tan buen papel, cuando menos estos resultados son alentadores.

El reconocimiento Baldrige no es un fin en sí mismo, sino más bien el medio para un fin. El proceso de preparación y postulación obliga a la compañía a examinar de cerca sus esfuerzos de calidad y proporciona una copia heliográfica para autoevaluarse. La aplicación Baldrige es un medio excelente para dilucidar problemas y sugerir dónde se necesita hacer mejoras.

ISO 9000

En 1992 la Organización Internacional de Estándares (ISO), con base en Suiza, estableció los lineamientos de ISO 9000. Este fue el primer intento para desarrollar un estándar uniforme verdaderamente global para la calidad. Para que una compañía obtenga el registro de ISO 9000, debe documentar cuidadosamente sus sistemas y sus procedimientos. Las compras, el manejo de los materiales, la manufactura y la distribución se sujetan todos a la documentación y la certificación de ISO. La certificación de ISO es muy diferente del proceso del premio Baldrige. La certificación de ISO no es un reconocimiento, ni siquiera es necesariamente un juicio sobre la calidad de los productos. Es una certificación de los procesos de manufactura y de negocios usados por la compañía.

Tras el establecimiento del estándar ISO vino una horda de organizaciones consultoras que se especializan en llevar a la compañía a través del proceso de certificación. La certificación no es barata. Además del costo directo de los consultores, hay costos indirectos importantes asociados con preparar la documentación necesaria. De acuerdo con John Rudin, presidente de la Compañía de abastecimientos Reynolds Aluminum:

Desde el punto de vista del desembolso, simplemente lograr el registro de una instalación se sitúa dentro del rango de $25 000 a $30 000. Para RASCO como un todo, usted está hablando de un millón o algo así, si fuéramos a certificar a todas las instalaciones. En este momento no vemos el valor de tener una placa que cuelgue en la pared por un millón de dólares, pero ciertamente vemos el valor de tener un proceso bueno, que regrese a lo básico, bien documentado, que es de lo que realmente se trata ISO de cualquier manera (Kuster, 1995).

La compañía que busca la certificación debe documentar procedimientos relacionados con la calidad, políticas relacionadas con el sistema y la contratación, y llevar registros. Según Velury (1995), una compañía debe esperar documentar aproximadamente 20 procedimientos para cada sección requerida para la certificación. Dado el gasto en tiempo y en dinero, ¿por qué debería molestarse alguien? Hay varias ventajas de obtener la certificación. Una proviene del proceso mismo. La documentación cuidadosa de las prácticas de calidad revela dónde fallan estas prácticas. El proceso de mejora continua comienza con el conocimiento de dónde está uno

parado. Además, actualmente muchas compañías requieren proveedores para obtener la certificación. Podría ser un prerrequisito para hacer negocios en muchas circunstancias. De acuerdo con Thomas Boldlund, presidente de Oregon Steel:

> ISO 9000 nos permite participar en mercados a los que hace 10 años no podíamos servir aunque quisiéramos. Cuando conocemos a un cliente, lo primero que pregunta es: "¿Puede usted cumplir los estándares de calidad de su competidor?" y segundo: "¿Puede usted producir los grados que actualmente producen los europeos y los japoneses?" Con ISO 9000, la respuesta es sí (Kuster, 1995).

El ISO continúa desarrollando estándares internacionales. En 1996 la organización anunció la serie ISO 14000 de estándares ambientales (Alexander, 1996). ISO 14000 es una serie de estándares internacionales para los sistemas de manejo ambiental. La motivación de ISO en la adopción de un estándar ambiental uniforme es la disparidad de estos estándares en diferentes países. Se espera que este nuevo estándar no solamente estimule el comercio internacional, sino lo que es más importante, mejore la calidad de la vida al cuidar el ambiente. Ambos lineamientos de ISO 9000 e ISO 14000 son los primeros pasos importantes en el movimiento hacia un estándar global uniforme.

Calidad: La línea inferior

¿Qué es lo que finalmente determina si un programa de calidad tiene éxito? Es la "línea inferior". Ya sea que el programa sea TQM, mejora continua, o seis sigma, la dirección de la compañía debe creer que la implementación finalmente va a conducir a mayores ganancias. El pensamiento tradicional es que las iniciativas de calidad son caras. Para alcanzar una alta confiabilidad y el apego a las especificaciones, podrían requerirse gastos sustanciales de capital (véase la figura 11-1). Solamente si el gasto se traduce a una línea inferior mejorada, ello lo justificará.

Que la calidad cueste más no es del todo claro. Además de erosionar la base de clientes, hay numerosos costos directos por la mala calidad. Estos incluyen los costos de los inventarios adicionales del trabajo en proceso, los costos de la chatarra, los costos del reprocesado, y los costos de inspección. Las compañías encaran riesgos crecientes de que los problemas de diseño y/o confiabilidad conducirán a demandas y acuerdos costosos.

El estudio de Garvin sobre la industria del aire acondicionado doméstico de 1981 a 1982 (Garvin, 1988), resumido en la tabla 11-7, muestra el dramático incremento de los costos de la calidad a medida que bajan la calidad del producto y la confiabilidad.

Aunque no es obvio en esta tabla, los fabricantes japoneses como grupo tuvieron tasas de defectos mucho más bajas que los mejores fabricantes de Estados Unidos. (Esto es más claro en el estudio de Garvin de 1983). Antes de analizar los resultados, debemos señalar algunas inconsistencias. Primero, el periodo de garantía en Estados Unidos es de cinco años y en Japón es de tres años. Esto tendería a sesgar los resultados a favor de los japoneses, aunque el sesgo es pequeño porque la mayoría de los costos por garantía se incurren en el primer año. Segundo,

TABLA 11-7 Calidad y costos de calidad en la industria de aire acondicionado doméstico

Agrupamiento de las compañías de acuerdo al rendimiento en la calidad	Costos de garantía promedio como porcentaje de las ventas	Costo total de la calidad (compañías japonesas), costos totales de las descomposturas (compañías estadounidenses) como un porcentaje de las ventas
Fabricantes japoneses	0.6%	1.3%
Las mejores plantas de Estados Unidos	1.8	2.8
Plantas buenas de Estados Unidos	2.4	3.4
Plantas regulares de Estados Unidos	2.7	3.9
Plantas malas de Estados Unidos	5.2	>5.8

Fuente: Garvin, 1988, p. 82.

los porcentajes en la tercera columna incluyen los costos de prevención, inspección, reprocesado, chatarra, y garantías para las compañías japonesas, y solamente reprocesado, chatarra y garantías para las compañías estadounidenses. Así, los porcentajes comparables en la tercera columna para las compañías estadounidense son aun más altos que aquellos reportados en la tabla.

Lo que es claro de este estudio es que los costos de la baja calidad son importantes. De hecho, se observó que eran hasta cinco veces más altos para los fabricantes malos de Estados Unidos que para los fabricantes japoneses. Sin embargo, esto no prueba de manera concluyente que sea necesariamente menos caro producir productos de calidad más alta. Estos datos no muestran las inversiones adicionales en equipo de capital, procesos, o el control estadístico de la calidad hechas por los fabricantes buenos, y por esa razón no son prueba de que la "calidad sea gratis". Sin embargo, otros estudios (véase en especial Schoeffler *et al.*, 1974) muestran que las compañías que producen productos de más alta calidad tienen una retribución más alta, una mayor participación en el mercado, y más ganancias que las compañías que producen productos de menor calidad.

11.14 DISEÑO DE LA CALIDAD EN EL PRODUCTO

Los métodos tradicionales de control de calidad se basan en el muestreo y la inspección. Los planes de muestreo y las políticas de inspección tienen el objetivo final de producir un porcentaje aceptable de artículos defectuosos. El considerar la calidad en términos del ciclo completo del producto, incluyendo el diseño, la producción, y el consumo, muestra que es económico diseñar la calidad en el producto. El especialista en estadística y consultor japonés Genichi Taguchi ha implementado un desarrollo reciente, que aplica el diseño estadístico de los experimentos al problema del diseño de producto. Las contribuciones iniciales de Taguchi fueron en los métodos de control de calidad *fuera de línea*. Su último trabajo incorporó también aspectos económicos.

Considere el siguiente ejemplo sencillo de Taguchi *et al.* (1989). Es bien sabido que, antes de la Segunda Guerra Mundial, la calidad de los bienes manufacturados en Japón era mala. En ese tiempo los japoneses buscaban competir en precio más que en calidad. Considere un producto japonés que se vendía a la mitad del precio de su competidor estadounidense. Si se basara una decisión de compra solamente en el precio, se habría escogido el producto japonés inferior. Sin embargo, al considerar el consumo debemos tomar en cuenta las pérdidas incurridas por el cliente al usar ese producto. Suponga que la pérdida por usar el producto hecho en Estados Unidos fuera igual al precio de compra, digamos P. Además, suponga que la paridad del producto hecho en Japón fuera nueve veces su precio de compra, por lo que suponemos que es $0.5P$. Entonces el costo total para el cliente del producto hecho en Estados Unidos habría sido $P + P = 2P$ y el costo total para el cliente del producto hecho en Japón habría sido $0.5P + (9)(0.5P) = 5P$. Así, el producto hecho en Japón habría costado al consumidor dos y media veces tanto como el hecho en Estados Unidos. Dada la experiencia con ambos productos, el consumidor finalmente se daría cuenta de que sus costos totales serían mayores con el producto menos caro.

Es precisamente este tipo de fenómeno el que ha conducido a máximas tales como "usted obtiene lo que usted paga", y explica en términos económicos simples por qué los consumidores deciden pagar más por la calidad.[4] La ironía de este ejemplo es con qué efectividad los japoneses han logrado colocarse en el otro lado de la ecuación. Explica por qué los automóviles japoneses continuaron vendiéndose bien en Estados Unidos aun cuando el tipo de cambio entre el dólar y el yen era tan desfavorable para los japoneses.

---

[4] Sin embargo, la afirmación inversa de que los productos más caros son necesariamente superiores no siempre es verdad. Véase Garvin (1988), p. 70, para un análisis de la correlación entre la calidad y el precio.

La calidad tiene un valor económico significativo para el consumidor, y el diseño de producto juega un papel importante en la calidad del producto. ¿Qué significa diseñar para la calidad? Significa que deberá minimizarse el número de partes que fallan fácilmente, o aquellas que complican significativamente al proceso de manufactura. En especial, ¿cómo puede simplificarse el diseño para eliminar partes pequeñas como los tornillos y los broches que son difíciles de ensamblar y que posiblemente sean puntos de conflicto posteriores? Un ejemplo de un diseño en el cual el número de partes se reduce a un mínimo es la Proprinter IBM. IBM desarrolló un video impresionante mostrando con qué facilidad puede armarse el producto a mano. Este diseño simple fue un factor importante en la confiabilidad y el éxito del producto en el mercado.

El ciclo de diseño es una parte importante de la cadena de la calidad. Taguchi *et al.* (1989) recomiendan los siguientes tres pasos en el ciclo de diseño de ingeniería:

1. *Diseño del sistema*. Este es el diseño básico del prototipo que cumple con las especificaciones de desempeño y tolerancia del circuito. Incluye la selección de materiales, partes, componentes y el ensamblado del sistema.
2. *Diseño de parámetros*. Después que se desarrolla el diseño del sistema, el siguiente paso es la optimización de los parámetros del sistema. Dado un diseño de sistema, generalmente hay varios parámetros del sistema cuyos valores deben determinarse. Un diseño de parámetros puede ser la ganancia de un transistor que es parte de un circuito. Es necesario encontrar una relación funcional entre el parámetro y la medida de desempeño del sistema para determinar un valor óptimo del parámetro. En el ejemplo, la medida del desempeño puede ser la salida de voltaje del circuito. La meta es encontrar el valor paramétrico que optimice la medida del desempeño. El método de Taguchi considera ese aspecto.
3. *Diseño de tolerancias*. El propósito de este paso es determinar los rangos permisibles para los parámetros cuyos valores se optimizan en el paso 2. Alcanzar el valor óptimo de un parámetro puede ser muy caro, mientras que un valor subóptimo podría dar la calidad deseada a un costo menor. El paso del diseño de tolerancias requiere una evaluación explícita sobre los costos asociados con los valores de los parámetros del sistema.

Los mismos conceptos pueden aplicarse al diseño del proceso de producción una vez que el diseño de producto ha sido terminado. La fase del diseño del sistema corresponde al diseño del proceso real de manufactura. En la fase del diseño de parámetros se identifican medidas que afectan al proceso de manufactura. Los ejemplos típicos son la variación de la temperatura, la variación de la materia prima y la variación del voltaje de entrada. En la fase del diseño de tolerancias, se determinan rangos aceptables para los parámetros identificados en la fase 2.

El área de control de la calidad fuera de línea (en oposición al tema de este capítulo, que podría denominarse como el control de la calidad en línea) incluye técnicas para alcanzar estos tres objetivos de diseño. Los métodos de Taguchi, basados en la teoría del diseño de experimentos, dan nuevos enfoques para resolver estos problemas.

La optimización de un valor paramétrico no siempre puede ser la solución general óptima. En algunos casos, el rediseño del producto para que sea menos sensible al parámetro en cuestión puede ser más económico. Con esta intención, Kackar (1985) describe a un fabricante japonés de mosaico quien resolvió el problema de la sensibilidad a la temperatura de esta manera. Citando el artículo:

> Un fabricante japonés de mosaico de cerámica sabía en 1953 que es más costoso controlar las causas de las variaciones en la manufactura que hacer que el proceso sea insensible a estas variaciones. La compañía de mosaicos Ina sabía que una distribución desigual de la temperatura en el horno causaba la variación del tamaño de los mosaicos. Como la distribución desigual de la temperatura era una causa asignable de la variación, un enfoque de control de la calidad del proceso habría sido diseñar métodos para controlar la distribución de la temperatura. Este enfoque habría incrementado el costo

de manufactura. La compañía quería reducir la variación del tamaño sin incrementar el costo. Por lo tanto, en vez de controlar la distribución de la temperatura trataron de encontrar una formulación para el mosaico que redujera el efecto de la distribución desigual de la temperatura sobre la uniformidad de los mosaicos. A través de un experimento diseñado, la compañía de mosaicos Ina encontró un método efectivo en costo para reducir la variación del tamaño del mosaico causada por la distribución desigual de la temperatura en el horno. La compañía encontró que un incremento del contenido de cal en la formulación del mosaico de 1 por ciento a 5 por ciento reducía la variación del tamaño del mosaico por un factor de 10. Este descubrimiento fue un avance importante para la industria del mosaico de cerámica.

El método de Taguchi se basa en suponer una función de pérdida, digamos $L(y)$, donde y es el valor de alguna característica funcional y $L(y)$ es la pérdida de la calidad medida en dólares. Para estar al nivel de gran parte de la teoría clásica de la estadística y del control, Taguchi recomienda una función cuadrática de pérdida. La forma cuadrática es el resultado de usar los dos primeros términos de un desarrollo en serie de Taylor. Dada una forma explícita para la función de pérdida, pueden encararse interrogantes como los beneficios de reducir las tolerancias y el valor de las diferentes políticas de inspección. El lector interesado puede consultar a Taguchi *et al.* (1989) y Logothetis y Wynn (1989) para ver un análisis sobre la teoría general. Dehnad (1989) trata las aplicaciones a industrias específicas.

Diseño, manufactura y calidad

La calidad comienza con el diseño del producto y con la manera en que se integra el diseño en el proceso de manufactura. El diseño más creativo del mundo es inútil si no puede fabricarse en forma económica hasta llegar a un producto confiable. La vinculación exitosa de los procesos de diseño y de manufactura es un signo distintivo del éxito del Japón en los productos para el consumidor. El movimiento del **diseño para la facilidad de manufactura** (DFM) en Estados Unidos tuvo sus raíces en los métodos japoneses de manufactura.

Boothroyd y Dewhurst (1989) y Boothroyd *et al.* (1994) fueron de los primeros en desarrollar un sistema de puntuación efectivo para diseños en términos de facilidad de manufacturabilidad. Los dos libros resumen la metodología desarrollada por estos dos ingenieros para un periodo de varios años. El primer libro se aboca a la eficiencia del ensamblado. La eficiencia del ensamblado es la relación del tiempo de ensamblado mínimo teórico entre una estimación del tiempo real de ensamblado con base en el diseño actual del producto. Este último libro trata aspectos más generales de DFM incluyendo el número de partes, los tipos de las partes, y los tipos de sujetadores. Estas reglas, que son muy detalladas, recomiendan preferir los diseños más simples con menos partes. Estos diseños conducen a productos que son más fáciles y menos caros de fabricar, y es menos probable que fallen con el uso.

Ulrich y Eppinger (1995) recomiendan que los diseñadores rastreen la complejidad del diseño vía un enfoque de marcador (véase el ejemplo 11.10). Un marcador proporciona una manera de comparar objetivamente diferentes diseños y un medio de rastrear la complejidad del proceso de manufactura para cada diseño de producto.

Ejemplo *11.10*

Marcador del ejemplo de la complejidad de la manufactura

Promotores de complejidad	Revisión 1	Revisión 2
Número de nuevas partes introducidas	6	5
Número de nuevos proveedores introducidos	3	2
Número de partes sobre pedido introducidas	2	3
Número de nuevas "herramientas principales" introducidas	2	2
Número de nuevos procesos de producción introducidos	0	0
Total	13	12

Fuente: Ulrich y Eppinger, 1995.

Este ejemplo es solamente para propósitos ilustrativos. En la práctica, el equipo tendría que decidir sobre la importancia relativa de los impulsores y aplicar pesos adecuados. Los marcadores como éstos obligan al equipo de diseño a examinar muy de cerca las consecuencias de sus decisiones sobre la manufactura.

Un ejemplo de un esfuerzo exitoso DFM fue el Proprinter de IBM. El Proprinter fue una impresora de matriz de punto centrada en el mercado de impresoras para PC que dominaron los japoneses en los inicios de la década de los 80 y siempre está en expansión. IBM desarrolló un video que muestra a alguien ensamblando la impresora a mano en cuestión de minutos. Al diseñar la Proprinter, IBM siguió la metodología clásica DFM. La Proprinter tenía muy pocas partes separadas y virtualmente no tenía tornillos ni sujetadores, sin que ello comprometiera su funcionalidad. El resultado fue que IBM pudo ensamblar la Proprinter en Estados Unidos y seguir siendo competitiva en el costo con los rivales japoneses (Epson, especialmente) que dominaban el mercado en ese momento. La Proprinter fue un producto muy exitoso de la IBM.

El número de partes en un producto no es la única medida de manufacturabilidad. La forma exacta del diseño de las partes y de su ensamblado también juega un papel importante. De acuerdo con Boothroyd y Dewhurst (1989), las características ideales de una parte son

- La parte debe insertarse en la parte superior del ensamblado.
- La parte es autoalineable.
- La parte no necesita orientarse.
- La parte requiere solamente una mano para el ensamblado.
- La parte no necesita herramientas.
- La parte se ensambla siguiendo un solo movimiento lineal.
- La parte se asegura inmediatamente con la inserción.

Aun cuando hay muchos éxitos claros sobre aplicar DFM, todavía falta que la metodología gane aceptación universal. De acuerdo con Boothroyd *et al.* (1994), las siguientes razones son las más comunes para no implementar DFM en la fase de diseño:

1. *No hay tiempo*. Los diseñadores están presionados por terminar sus diseños rápidamente para minimizar el tiempo entre diseño y manufactura para un nuevo producto. El enfoque DFM es de tiempo intensivo. El diseño para reducir los costos de ensamblado y la complejidad del producto no puede hacerse al azar.
2. *Aquí no se inventa*. Las ideas nuevas siempre son rechazadas. Sería mejor si el ímpetu de DFM proviniera de los mismos diseñadores, pero con mayor frecuencia proviene de la gerencia. Los diseñadores resienten que los extraños les arrojen un nuevo enfoque (como lo hace cualquiera).
3. *Costos bajos de ensamblado*. Como frecuentemente los costos de ensamblado toman en cuenta una pequeña porción de los costos totales de manufactura, podría argumentarse que no tiene mucho caso hacer un diseño para el análisis de ensamblado (DFA). Sin embargo, frecuentemente los ahorros pueden ser mayores de lo que se pensaría.
4. *Bajo volumen*. Podría argumentarse que el análisis DFM no vale la pena para los artículos de bajo volumen. Boothroyd *et al.* (1994) argumentan que lo opuesto es verdad. Si los volúmenes son bajos, es poco probable el rediseño una vez que comienza la producción. Esto significa que el hacerlo bien la primera vez es aun más importante.
5. *Ya lo hicimos*. Muchas compañías han usado algunas reglas generales sencillas para el diseño (tal como limitar el número de dobleces en una parte de hoja laminada). Aun cuando estas reglas tienen sentido en forma aislada, es poco probable que conduzcan al mejor diseño general para el producto.

6. *DFM conduce a productos a los cuales es difícil darles servicio*. Posiblemente esto no es verdad. Los productos que son más fáciles de ensamblar son más fáciles de desensamblar, y por tanto es más fácil darles servicio.

Dvorak (1994) ofrece otras razones para la baja aceptación de DFM. Los sistemas contables clásicos tal vez no puedan reconocer los ahorros en costo de los nuevos diseños. Un diseño que reduce los costos fijos de instalación no sería visto como efectivo en costo ya que en muchos sistemas contables los costos fijos se consideran parte de los costos administrativos. Un sistema contable basado en la actividad no tendría este problema. Sin embargo, la mayoría concuerda en que el mayor obstáculo para que se acepte DFM es la resistencia al cambio.

Aun cuando el movimiento DFM tal vez no gane aceptación con la rapidez que les gustaría a algunos, es evidente que hay una conciencia creciente en torno al uso de estos poderosos métodos. Dvorak (1994) discute varias historias de éxito de la implementación de DFM. Una es la de Coors, donde se mejoraron el rendimiento de la producción, la calidad, y la confiabilidad de la entrega. De acuerdo con Dvorak:

> Está cundiendo la noticia en la industria de cómo DFM puede traer éxitos similares a los de Coors. Muchas organizaciones están enarbolando la bandera DFM. El Departamento de Comercio de Estados Unidos, entre otros, ha iniciado proyectos DFM en seis de sus centros regionales de manufactura. General Motors ha dependido de DFM para energizar sus esfuerzos concurrentes de ingeniería con una respuesta asombrosa: está ahorrando el 20 por ciento del costo total del automóvil para los modelos 1992 a los cuales se les aplicó. DFM y los conceptos de diseño para el ensamblado ya han hecho maravillas en el recorte de material y en los costos de manufactura para una gama de productos. No hay nuevas ideas que estén infiltrando a la disciplina.

La competencia internacional continúa calentándose. Los japoneses en especial han demostrado que puede tenerse éxito mediante un análisis cuidadoso y razonado durante la fase de diseño y desarrollo del producto. Han sido adeptos al diseño concurrente del producto y del proceso. El resultado son productos que pueden manufacturarse con mayor eficiencia y que pueden trabajar mejor que los de los competidores. Aun cuando DFM es sólo una parte del pastel, provee una metodología para vincular el diseño y la manufactura.

En última instancia, no debemos olvidar que el diseño de producto se prolonga mucho más allá de solamente los aspectos de la manufacturabilidad. Los aspectos estéticos también son importantes, y pueden ser el factor dominante para algunos productos. Otro aspecto importante es el proceso de angostamiento del campo de las opciones hasta un diseño final. Ya que el interés de este libro radica en los aspectos relacionados con la manufactura, no discutiremos aquí estos aspectos de diseño más amplios, el lector interesado puede consultar a Pugh (1991).

11.15 NOTAS HISTÓRICAS

El deseo de mantener la calidad de los bienes manufacturados no tiene nada de nuevo. El hombre prehistórico cuyas armas y herramientas no funcionaran, no sobrevivía. Sin embargo, los métodos de control estadístico de la calidad estudiados en este capítulo se diseñaron apenas en los últimos 70 años o algo así. Walter Shewhart, quien era un empleado de los laboratorios Bell Telephone, ideó la gráfica de control. Su monografía de 1931 (Shewhart, 1931) resumió sus contribuciones en esta área. H. F. Dodge y H. G. Romig, también empleados de los laboratorios Bell, generalmente reciben el crédito de haber establecido los fundamentos de la teoría del muestreo de aceptación.

Al igual que con la mayor parte de la nueva metodología, la industria estadounidense demoró en adoptar las técnicas de control de la calidad. Sin embargo, como parte del esfuerzo de la guerra, el gobierno de Estados Unidos decidió adoptar los métodos de inspección de

muestreo para los pertrechos militares en 1942. Como resultado de éstas y otras actividades gubernamentales en tiempos de guerra, el conocimiento y la aceptación de las técnicas estudiadas en este capítulo se popularizaron después de la guerra. De hecho, los planes de muestreo frecuentemente se describen usando términos del ejército de Estados Unidos. Muchas de las técnicas de muestreo de aceptación estudiadas en este capítulo son parte de la Norma Militar 105. Esta última y sus diferentes revisiones y adiciones forman la base de la mayoría de los planes de muestreo de aceptación que se usan actualmente.

Abraham Wald fue responsable del desarrollo de la teoría del análisis secuencial, que forma la base de las fórmulas que aparecen en la sección 11.11. El trabajo de Wald demostró ser un hito importante para el avance de la teoría del muestreo de aceptación. Wald era parte de un equipo de investigación organizado en la Universidad de Columbia en 1942. El gobierno de Estados Unidos consideró que su trabajo era tan importante que detuvo la publicación hasta junio de 1945.

W. Edwards Deming, quien visitó Japón al inicio de la década de los 50 para dictar una serie de conferencias sobre los métodos de control de calidad, recibe gran parte del crédito por la transferencia de la tecnología del control estadístico de la calidad a Japón. Actualmente, el muy prestigiado premio Deming de control de la calidad, establecido por Deming en 1951, sigue siendo un símbolo de la dedicación japonesa a la misma.

La mayoría de los métodos estudiados en este capítulo se usan ampliamente en la industria de Estados Unidos y en el extranjero. Sin embargo, los modelos de optimización para diseñar las gráficas de control no gozan el mismo nivel de aceptación. El modelo esbozado en la sección 11.7 es de Baker (1971), aunque anteriormente se desarrollaron modelos económicos más complejos y exhaustivos para diseñar gráficas X (vea, por ejemplo, Duncan, 1956). Para un estudio exhaustivo de la historia del movimiento de la calidad en Estados Unidos, vea Kolesar (1993).

11.16 Resumen

En este capítulo describimos las técnicas del diseño de las gráficas de control y del muestreo de aceptación. La base de la metodología de este capítulo es la probabilidad y la estadística clásicas. Los puntales de las gráficas de control incluye resultados fundamentales para la teoría de la probabilidad como la ley de los grandes números y el teorema del límite central.

Estudiamos las *gráficas de control* en la primera mitad de este capítulo. Una gráfica de control es un dispositivo gráfico que se usa para determinar cuándo ha ocurrido una desviación en el valor de un parámetro de la distribución básica de alguna cantidad medible. Cuando esto ocurre, se dice que el proceso se ha salido del control. El diseño de la gráfica de control depende del parámetro específico que se está monitoreando.

La gráfica de control más común es la gráfica $\overline{X}$. Una gráfica $\overline{X}$ se diseña para monitorear una variable medible individual, como el peso o la longitud. Los subgrupos de tamaño n se muestrean sobre una base regular y la media muestral del subgrupo se calcula y se coloca en una gráfica. Como la media muestral tiene una distribución normal aproximada independientemente de la forma en que la población se distribuya, la posibilidad de que una observación individual se sitúe fuera de los límites tres sigma es tan pequeña que cuando ocurre un evento de este tipo es poco probable que se deba al azar. Más bien, es probable que sea el resultado de una desviación de la media verdadera del proceso.

El segundo tipo de gráfica de control que se trata en este capítulo es la gráfica R. Una *gráfica* R se diseña para determinar si la variación del proceso se ha desviado. El símbolo R significa rango. El rango es la diferencia entre las observaciones más grande y más pequeña en un subgrupo. Debido a que existe una estrecha relación entre el valor del rango de la muestra y la varianza intrínseca de la población, las gráficas R se usan para medir la estabilidad de la varianza de un proceso.

Con frecuencia las gráficas de control para variables no son apropiadas. La *gráfica p* es una gráfica de control para los atributos. Cuando se usan gráficas p, los artículos se clasifican como aceptables o no. La gráfica p utiliza la aproximación normal de la distribución binomial y puede usarse cuando los subgrupos son de tamaño igual o variable.

La última gráfica de control presentada fue la *gráfica c*. La gráfica c se usa para monitorear el número de artículos defectuosos en una unidad de producción. El parámetro c es el número promedio o tasa de defectos por unidad de producción. La gráfica c se basa en la distribución de Poisson, y los límites de control se obtienen usando la aproximación normal a la de Poisson.

En la sección 11.6 se consideró un modelo para el diseño económico de las gráficas $\overline{X}$. Las variables de decisión para el modelo fueron el tamaño de cada subgrupo n de la muestra, y el número k de desviaciones estándar que se usan para señalar una condición fuera de control. El modelo se basa en la hipótesis de que el proceso se sale de control en forma aleatoria con una probabilidad conocida π. También se supone que se conocen el costo del muestreo, el costo de la búsqueda de una causa asignable cuando ocurre una señal de fuera de control, y el costo de operación del sistema para una condición fuera de control.

El capítulo también consideró el *muestreo de aceptación*. El propósito de un esquema de muestreo de aceptación es determinar si la proporción de artículos defectuosos en un lote grande de artículos es aceptable con base en los resultados del muestreo de un número relativamente pequeño de artículos provenientes del lote. El plan de muestreo más sencillo es el muestreo simple. Un *plan de muestreo simple* está especificado por dos números: n y c. Interprete n como el tamaño de la muestra y c como el nivel de aceptación. Un *plan de muestreo doble* requiere que se especifiquen cinco números: n_1, n_2, c_1, c_2 y c_3, aunque con frecuencia $c_2 = c_3$. Con base en los resultados de un tamaño de muestra inicial de n_1, el lote se acepta o se rechaza, o se extrae una muestra adicional de tamaño n_2. La ampliación lógica del muestreo doble es el *muestreo secuencial*, en el cual los artículos se muestran uno a la vez y se toma una decisión después de muestrear cada artículo para decidir si se acepta el lote, se rechaza el lote, o se continúa el muestreo.

La *administración de la calidad total* es un término que oímos más frecuentemente ya que las compañías de Estados Unidos luchan para competir con sus competidores europeos y japoneses. Estudiamos las ocho dimensiones de la calidad de Garvin: (1) desempeño, (2) características,(3) confiabilidad,(4) conformidad, (5) durabilidad, (6) capacidad de servicio, (7) estética y (8) calidad percibida. También se discutieron métodos para atender la voz de los clientes, como el *análisis conjunto* y el *despliegue de la función de calidad* (QFD). Con objeto de que tenga éxito el TQM, la actividad de la calidad debe trascender los límites funcionales y departamentales. Un enfoque que intentó hacer esto fue los *círculos de calidad*. El programa requirió una inversión mínima y que se reestructurara la parte de la gerencia, y como resultado no ha tenido mucho éxito. El benchmarking proporciona un medio para que una compañía compare su desempeño con sus competidores y aprenda las "mejores prácticas" de la industria. Dos premios nacionales muy codiciados son el premio Deming en Japón y el reconocimiento Baldridge en Estados Unidos. Estas distinciones reconocen los esfuerzos excepcionales de la industria para implementar la calidad.

Los *métodos de calidad fuera de línea* se dirigen al problema de diseñar la calidad dentro del producto. Los métodos Taguchi, basados principalmente en la teoría del diseño de experimentos, son un desarrollo importante en esta área. Los métodos Taguchi identifican procesos importantes y parámetros de diseño e intentan encontrar valores generales óptimos para estos parámetros, relacionados con alguna medición del desempeño del sistema. El capítulo concluyó con una discusión en torno al diseño para manufacturabilidad y las contribuciones de Boothroyd y Dewhurst a esta área.

Problemas adicionales sobre calidad y aseguramiento

49. ¿De qué maneras podrían contribuir a la mala calidad cada uno de los factores listados?

a) Administración
b) Mano de obra
c) Mantenimiento del equipo
d) Diseño del equipo
e) Control y monitoreo
f) Diseño de producto

50. La figura 11-1 presenta una ilustración conceptual del intercambio entre el costo del proceso y los costos de las pérdidas debidas a la mala calidad. ¿Cuáles son los costos de la mala calidad y qué dificultades podrían surgir al intentar medir estos costos?

51. Las gráficas $\overline{X}$ y R se mantienen en una dimensión individual de la calidad. Ocurre una desviación repentina en el proceso, haciendo que la media del proceso se incremente en 2σ, donde σ es la desviación estándar verdadera del proceso. No ocurren desviaciones en la variación del proceso. Suponiendo que la gráfica $\overline{X}$ se basa en los límites 3σ y en los subgrupos de tamaño $n = 6$, ¿qué proporción de los puntos en la gráfica $\overline{X}$ esperaría usted que se situaran fuera de los límites después que ocurra la desviación?

52. XYZ produce rodamientos para ruedas de bicicleta y monitorea el proceso con una gráfica $\overline{X}$ del diámetro de los rodamientos. La gráfica $\overline{X}$ se basa en subgrupos de tamaño 4. El valor meta es 0.37 pulgadas, y los límites superior e inferior son 0.35 y 0.39 pulgadas, respectivamente (suponga que se basan en límites tres sigma). Wheeler, quien compra los rodamientos de XYZ para construir las ruedas, requiere tolerancias de 0.39 ± 0.035 pulgadas. Los rodamientos demasiado grandes se rebajan y los rodamientos demasiado pequeños se descartan. ¿Qué proporción de los rodamientos tiene que rebajar Wheeler y que proporción debe descartar? ¿Puede sugerir lo que debe hacer Wheeler para reducir la proporción de rodamientos que debe descartar?

53. Un proceso que está dentro del control estadístico tiene un valor estimado de la media de 180 y una desviación estándar estimada de 26.

a) Con base en subgrupos de tamaño 4, ¿cuáles son los límites de control para las gráficas $\overline{X}$ y R?
b) Suponga que ocurre una desviación de la media de modo que el nuevo valor de la media es 162. ¿Cuál es la probabilidad de que se detecte la desviación en el primer subgrupo después que ocurre la desviación?
c) En promedio, ¿cuántos subgrupos tendrían que ser muestreados antes de que ocurriera la desviación, para descubrirla?

54. Considere los datos presentados en la tabla 11-1 para el ejemplo del brazo de lectura y escritura. Suponga que se construye una gráfica R basada solamente en los números de muestra 1 a 15.

a) ¿Cuál es el estimado de σ obtenido solamente de estas 15 observaciones?
b) ¿Cuáles son los valores de LSC y LIC para una gráfica R basada solamente en estas observaciones?

55. Discuta las ventajas y desventajas de las siguientes estrategias en el diseño de la gráfica de control. En especial, ¿cuáles son los intercambios económicos pertinentes a cada estrategia?

a) Seleccionar un valor muy pequeño de α.
b) Seleccionar un valor muy pequeño de β.

c) Seleccionar un valor grande de n.

d) Seleccionar un valor pequeño para el intervalo de muestreo.

56. La construcción de la gráfica p descrita en este capítulo requiere la especificación de ambos niveles, LSC y LIC. ¿Es significativo el LIC en este contexto? En especial, ¿qué significa si un valor observado de p es menor que el LIC? Si solamente se usa el LSC para señalar una condición fuera de control, ¿deberá modificarse de alguna manera el cálculo del LSC? (Sugerencia: la definición de la probabilidad del error Tipo 1, α, será diferente. La prueba de hipótesis para determinar si el proceso está bajo control es de una cola en lugar de dos colas.)

57. Un fabricante de aparatos grandes mantiene una gráfica p para la producción de máquinas lavadoras. Las máquinas pueden rechazarse debido a defectos cosméticos o funcionales. Con base en el muestreo de 30 máquinas cada día durante 50 días consecutivos, el estimado actual de p es .0855.

a) ¿Cuáles son los límites de control para una gráfica p? (Suponga límites 3σ.)

b) Suponga que el porcentaje de máquinas lavadoras defectuosas aumenta a 20 por ciento. ¿Cuál es la probabilidad de que se detecte esta desviación en el primer día después de ocurrido?

c) En promedio, ¿cuántos días se requerirían para detectar la desviación?

58. Un fabricante de computadoras personales, No Name, compra chips DRAM de 64K de dos fabricantes diferentes, A y B. No Name usa el siguiente plan de muestreo: se extrae una muestra de 10 por ciento de los chips y se rechaza el lote si se descubren dos o más chips defectuosos. Los dos fabricantes suministran los chips en lotes de 100 y 1 000, respectivamente.

a) Para cada fabricante, determine la proporción verdadera de artículos defectuosos en el lote que conduciría a la aceptación del 90 por ciento de los lotes. Usted puede usar la aproximación de Poisson para sus cálculos.

b) ¿Diría usted que este plan es justo?

59. Grafique las curvas AOQ para los fabricantes A y B mencionados en el problema 58. Estime los valores de AOQL en cada caso.

60. Considere el plan de muestreo discutido en el problema 58. ¿Sería mejor un plan que rechazara el lote si más del 10 por ciento de los chips en una muestra fueran defectuosos? ¿Cuál de los dos fabricantes mencionados llevaría ventaja si se adoptara este plan?

61. Suponiendo AQL $=$ 5 por ciento y LTPD $=$ 10 por ciento, determine los valores de α y β para el plan descrito en el problema 58 para los fabricantes A y B, y para el plan descrito en el problema 60 para los fabricantes A y B.

62. Grafique las curvas OC para el plan de muestreo descrito en el problema 58 para ambos fabricantes A y B.

63. Se usan gráficas de control $\overline{X}$ para mantener el control de la manufactura de las cajas usadas para alojar una marca genérica de computadora personal. Se mantienen gráficas separadas para la longitud, el ancho y la altura. La gráfica de longitud tiene LSC $=$ 20.5 pulgadas, LIC $=$ 19.5 pulgadas, y un valor meta de 20 pulgadas. Esta gráfica se basa en el uso de subgrupos de tamaño 4 y en límites tres sigma. Sin embargo, las especificaciones del cliente requieren que la longitud meta debe ser de 19.75 pulgadas con una tolerancia de ± 0.75 pulgadas. ¿Qué porcentaje de las cajas fletadas estará situado fuera de las especificaciones del cliente?

64. Se usa una gráfica p para monitorear la fracción defectuosa de un circuito integrado que va a usarse en un marcapasos comercial. Se toma una muestra de 15 circuitos de la producción de cada día durante 30 días hábiles consecutivos. Se descubre un total de 17 artículos defectuosos durante este periodo.

    a) Determine los límites de control tres sigma para este proceso.

    b) Suponga que α, la probabilidad del error Tipo 1 (es decir, la probabilidad de concluir que el proceso está fuera de control cuando está bajo control), se establece como .05. ¿Qué límites de control obtiene ahora? (Use una aproximación normal para sus cálculos.)

65. Un plan de muestreo simple se usa para determinar la aceptabilidad de los fletes de un ensamblado de rodamientos que se usa al fabricar patinetas. Para lotes de 500 rodamientos, se toman muestras de $n = 20$. El lote se rechaza si se encuentran en la muestra cualesquiera artículos defectuosos.

    a) Suponga que AQL = .01 y LTPD = .10. Encuentre α y β.

    b) ¿Este plan tiene más ventajas para el consumidor o para el productor?

66. Un plan de muestreo doble se construye como sigue. De un lote de 200 artículos, se extrae una muestra de 10 artículos. Si hay cero artículos defectuosos, se acepta el lote. Si hay dos o más artículos defectuosos, se rechaza el lote. Si hay exactamente un artículo defectuoso, se extrae una segunda muestra de 10 artículos. Si el número combinado de artículos defectuosos en ambas muestras es de dos o menos, se acepta el lote; de lo contrario se rechaza. Si el lote tiene 10 por ciento de artículos defectuosos, ¿cuál es la probabilidad de que se acepte?

67. Hammerhead produce clavos de servicio pesado, que son comprados por Modulo, un fabricante de casas prefabricadas. Modulo compra los clavos en lotes de 10000 y somete una muestra a pruebas destructivas para determinar la aceptabilidad del lote. Modulo ha establecido un AQL de 1 por ciento y un LTPD de 10 por ciento.

    a) Suponiendo un plan de muestreo simple con $n = 100$ y $C = 2$, encuentre α y β.

    b) Obtenga el plan de muestreo secuencial que alcance los mismos valores de α y β como el plan de muestreo simple obtenido en el inciso *a*.

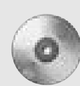
    c) Al estimar la curva ASN, encuentre el valor máximo del tamaño de muestra esperado que Modulo va a requerir si usa el plan secuencial obtenido en el inciso *b*.

    d) Suponga que se usa el plan de muestreo secuencial obtenido en el inciso *b*. Se prueban cien clavos con los siguientes resultados: los primeros 80 son aceptables, el 81 es defectuoso, y los 19 restantes son aceptables. Al graficar las regiones de aceptación y de rechazo, determine si el plan de muestreo secuencial obtenido en el inciso *b* recomendaría la aceptación o el rechazo antes de probar el clavo 100.

68. Para el plan de muestreo simple obtenido en el inciso *b* del problema 67, suponga que los lotes que no pasan se regresan a Hammerhead.

    a) Estime la gráfica de la curva AOQ calculando AOQ(p) para diferentes valores de p.

    b) Usando los resultados del inciso *a*, estime la proporción máxima de clavos defectuosos que Modulo estará usando en su construcción.



69. Veinte conjuntos de cuatro mediciones de los diámetros en pulgadas de pelotas de golf Hot Shot fueron

Muestra				
1	2.13	2.18	2.05	1.96
2	2.08	2.10	2.02	2.20
3	1.93	1.98	2.03	2.06
4	2.01	1.94	1.91	1.99
5	2.00	1.90	2.14	2.04
6	1.92	1.95	2.02	2.05
7	2.00	1.94	2.00	1.90
8	1.93	2.02	2.04	2.09
9	1.87	2.13	1.90	1.92
10	1.89	2.14	2.16	2.10
11	1.93	1.87	1.94	1.99
12	1.86	1.89	2.07	2.06
13	2.04	2.09	2.03	2.09
14	2.15	2.02	2.11	2.04
15	1.96	1.99	1.94	1.98
16	2.03	2.06	2.09	2.02
17	1.95	1.99	1.87	1.92
18	2.05	2.03	2.06	2.04
19	2.12	2.02	1.97	1.95
20	2.03	2.01	2.04	2.02

a) Ingrese los datos en una hoja de cálculo y calcule las medias y los rangos para cada muestra.

b) Usando los resultados de *a*, desarrolle gráficas $\overline{X}$ y R similares a las figuras 11-7 y 11-8. Suponga límites tres sigma para la gráfica $\overline{X}$.

c) Desarrolle un histograma basado en las 80 observaciones. Suponga intervalos de clase (1) 1.80–1.849, (2) 1.85–1.899, (3) 1.90–1.949, (4) 1.95–1.999, (5) 2.0–2.049, (6) 2.05–2.099, (7) 2.10–2.149, (8) 2.15–2.20. Con base en su histograma, ¿qué distribución describiría con exactitud el diámetro de una pelota de golf seleccionada aleatoriamente?



70. Una gráfica *p* se usa para monitorear el número de cortadoras de pasto montables producidas. Los números que se devuelven para el reprocesado porque no pasaron la inspección son

Día	Número producido	Número rechazado
1	400	23
2	480	18
3	475	24
4	525	34
5	455	17
6	385	17
7	372	12
8	358	19
9	395	24
10	405	29
11	385	16
12	376	19
13	395	23
14	405	14
15	415	25
16	440	34
17	380	26
18	318	19

Ingrese los datos en una hoja de cálculo y calcule los valores Z estandarizados para una gráfica p. Grafique los valores Z. ¿Está este proceso bajo control?



71. Se extraen muestras de tamaño 50 de lotes de 1 000 artículos. El lote se rechaza si hay más de dos artículos defectuosos en la muestra. Usando una aproximación binomial, grafique la curva OC como una función de p, la proporción de artículos defectuosos en el lote. Para un AQL de .01 y un LTPD de .10, encuentre α y β.

   a) Grafique la curva OC e identifique las probabilidades de los errores Tipo 1 y Tipo 2 (es decir, desarrolle una gráfica similar a la figura 11-14).

   b) Grafique la curva AOQ e identifique el valor de AOQL.



72. *a*) Desarrolle una hoja de cálculo a partir de la cual pueda obtenerse una gráfica como la de la figura 11-17 para el muestreo secuencial. Almacene los valores de p_0, p_1, α, y β en las localidades de la celda de modo que puedan alterarse a voluntad. Imprima una gráfica para $p_0 = .05$, $p_1 = .20$, $\alpha = .05$, y $\beta = .10$. Considere que $n \leq 100$.

   b) El muestreo secuencial condujo a lo siguiente: los primeros 40 artículos fueron buenos, el artículo 41 fue defectuoso, el artículo 68 fue defectuoso, y los artículos 86 y 87 fueron defectuosos. Coloque estos resultados en la gráfica que usted obtuvo en el inciso *a*. ¿El lote se acepta, se rechaza, o ninguna de las dos cosas durante o antes de probar el artículo 87?

Apéndice **11-A**

Aproximación de distribuciones

En este capítulo se usaron varias aproximaciones de probabilidad. En este apéndice se discutirá qué motiva y justifica estas aproximaciones.

La complejidad de una distribución de probabilidades depende del número de parámetros que se requieren para especificarla. Las distribuciones consideradas en este capítulo en orden descendiente de complejidad son

Distribución	Parámetros
1. Hipergeométrica	n, N, M
2. Binomial	n, p
3. Poisson	λ
4. Normal	μ, σ

En esta tabla no queda claro por qué la distribución normal debe ser más simple que la de Poisson, porque la normal es una distribución de dos parámetros y, la de Poisson, de uno. La razón es que todas las probabilidades normales pueden obtenerse a partir de una sola tabla de la distribución normal estándar, y la distribución de Poisson debe tabularse separadamente para valores diferentes de λ.

La *aproximación binomial a la hipergeométrica*. La distribución hipergeométrica (cuya fórmula aparece en la sección 11.9) es la probabilidad de extracción de n artículos de un lote de N artículos de los cuales M son defectuosos, y entonces hay exactamente m artículos defectuosos en la muestra. Puede pensarse que el experimento que da lugar a la hipergeométrica es el muestreo de artículos uno por uno sin remplazo. Si N es mucho mayor que n, la probabilidad de que cualquier artículo muestreado sea defectuoso es muy cercana a M/N. En ese caso, la probabilidad hipergeométrica se acercaría a la probabilidad binomial con $p = M/N$ y $n = n$.

Observe que la distribución binomial corresponde a un muestreo *con* remplazo. Si $N > 10n$, la binomial debe proporcionar una aproximación adecuada a la hipergeométrica.

La aproximación de Poisson a la binomial. La distribución de Poisson puede obtenerse como el límite de la binomial cuando $n \to \infty$ y $p \to 0$, pero con el producto np permaneciendo constante. Escriba $\lambda = np$. Entonces para n grande y p pequeña

$$p\{X = m\} \approx \frac{e^{-\lambda}\lambda^m}{m!} \qquad \text{para } m = 0, 1, 2, \ldots$$

No es obvio bajo qué circunstancias sea adecuada esta aproximación. En general, $p < .1$ y $n > 25$ deberían ser válidos, pero si p es muy pequeño, entonces son aceptables valores más pequeños de n, y si n es muy grande, entonces son aceptables valores mas grandes de p. Por ejemplo, para $n = 10$ y $p = .01$, la probabilidad binomial de que $X = 1$ es .0914 y la probabilidad de Poisson es .0905. Los valores de p cercanos a 1, tal como $p = .99$, también serían aceptables porque la distribución binomial con $p = .01$ es una imagen espejo de una distribución binomial con $p = .99$.

Aproximaciones normales. El teorema del límite central dice (a grandes rasgos) que la distribución de una suma de n variables aleatorias independientes con distribución idéntica se aproxima a la distribución normal cuando n se agranda. Como la distribución normal se obtiene sumando n variables aleatorias de Bernoulli independientes con distribución idéntica, si n es grande la normal da una buena aproximación a la binomial. Como la aproximación normal es más exacta cuado p se acerca a .5, una buena regla general es que la aproximación debe usarse solamente si $np(1 - p) > 5$.

Siempre que se use la distribución normal para aproximar cualquier otra distribución, es necesario expresar μ y σ en términos de los parámetros originales. En el caso binomial, $\mu = np$ y $\sigma = \sqrt{np(1 - p)}$.

Como la variable aleatoria normal es continua y la variable aleatoria binomial es discreta, la aproximación puede mejorarse usando la "corrección por continuidad". En el caso binomial, los eventos $\{X > 2\}$ y $\{X \geq 3\}$ son idénticos, pero en el caso normal no lo son. La corrección por continuidad sugeriría la aproximación de cada uno de estos casos por $\{X > 2.5\}$. La corrección por continuidad es expresar el evento original en términos de $>$ y $\geq$ (o $<$ y $\leq$) y usar un número truncado a la mitad entre los dos.

Por ejemplo, suponga que $n = 25$, $p = .40$, y deseamos determinar $P\{X \leq 10\}$. Como $\{X \leq 10\} = \{X < 11\}$, el truncamiento de la corrección por continuidad es 10.5. La probabilidad binomial exacta es .5858. La aproximación normal para 10 da

$$P\{X \leq 10\} \approx P\left\{Z < \frac{10 - (25)(.40)}{\sqrt{(25)(.40)(.60)}}\right\} = P\{Z < 0\} = 0.5,$$

y con la corrección por continuidad

$$P\{X \leq 10\} \approx P\left\{Z < \frac{10.5 - (25)(.40)}{\sqrt{(5)(.40)(.60)}}\right\} = P\{Z < .2041\} = 0.5948.$$

La distribución normal también puede usarse para aproximar a la de Poisson si λ es grande ($\lambda > 10$). En ese caso, use $\mu = \lambda$, $\sigma = \sqrt{\lambda}$, y la corrección por continuidad como se describe. Por ejemplo, suponga que deseamos usar una aproximación normal de la probabilidad de que una variable aleatoria de Poisson con parámetro $\lambda = 15$ sobrepase a 8. Como $\{X > 8\} = \{X \geq 9\}$, el truncamiento de la corrección por continuidad se sitúa en 8.5. Así,

$$P\{X > 8\} \approx P\left\{Z > \frac{8.5 - 15}{\sqrt{15}}\right\} = P\{Z > -1.68\} = .9535.$$

La probabilidad de Poisson exacta es .9626.

Apéndice 11-B

Glosario de notación para el capítulo 11 sobre aseguramiento de la calidad

Nota: Este capítulo usa la notación aceptada para las gráficas de control y el muestreo por aceptación. Como resultado, el mismo símbolo puede tener un significado en el contexto de las gráficas de control y otro en el contexto del muestreo de aceptación.

a_1 = Costo del muestreo de un artículo.

a_2 = Costo de buscar una causa asignable.

a_3 = Costo de operar el proceso en un estado fuera de control.

AOQ = Límite de calidad de salida promedio.

AOQL = Límite de calidad de salida promedio. El valor máximo de la curva AOQ.

AQL = Nivel aceptable de calidad.

α = $P\{$error Tipo 1$\}$.

Uso en la gráfica de control: α representa la probabilidad de obtener una señal fuera de control cuando el proceso está bajo control.
Uso en el muestreo de aceptación: α es la probabilidad de rechazar lotes buenos.

β = $P\{$error Tipo 2$\}$.

Uso en la gráfica de control: β representa la probabilidad de no obtener una señal fuera de control cuando el proceso se ha salido del control.
Uso del muestreo de aceptación: β es la probabilidad de aceptar lotes malos.

c = { *Uso en la gráfica de control:* el número esperado de defectos por unidad de producción
Uso en el muestreo de aceptación: el nivel de aceptación para un plan de muestreo simple. ·

c_1 = Nivel de aceptación para la primera muestra de acuerdo con un plan de muestreo doble.

c_2 = Nivel de rechazo para la primera muestra de acuerdo con un plan de muestreo doble.

c_3 = Nivel de aceptación para la primera y la segunda muestras de acuerdo con un plan de muestreo doble. (Frecuentemente $c_2 = c_3$).

d_2 = Una constante que depende de n, la cual relaciona a $\overline{R}$ con σ.

d_3 = Una constante que depende de n, la cual, cuando se multiplica por $\overline{R}$, da el límite inferior de control para una gráfica R.

d_4 = Una constante que depende de n, la cual, cuando se multiplica por $\overline{R}$, da el límite superior de control para una gráfica R.

δ = Magnitud supuesta de la desviación de la media tal como se mide en desviaciones estándar para el diseño económico de las gráficas $\overline{X}$.

LIC = Límite inferior de control de una gráfica de control.

LTPD = Tolerancia del lote en cuanto al porcentaje de artículos defectuosos. Nivel de calidad inaceptable.

M = Número de artículos defectuosos en un lote.

μ = Media poblacional.

N = Número de artículos en un lote. Se usa en el muestreo por aceptación.

$n = \begin{cases} \textit{Uso en la gráfica de control de calidad:} \text{ el tamaño de cada subgrupo para una gráfica } \overline{X}. \\ \textit{Uso en muestreo de aceptación:} \text{ número de artículos muestreados de un lote para un plan de muestreo simple.} \end{cases}$

n_1 = Tamaño de la primera muestra para un plan de muestreo doble.

n_2 = Tamaño de la segunda muestra en un plan de muestreo doble.

OC(p) = Curva de características de operación.

$p = \begin{cases} \textit{Uso en la gráfica de control:} \text{ proporción verdadera de los artículos defectuosos producidos.} \\ \textit{Uso en muestreo de aceptación:} \text{ proporción verdadera de artículos defectuosos en un lote.} \end{cases}$

p_0 = AQL.

p_1 = LTPD.

π = Probabilidad de que el proceso se salga de control en un periodo simple.

R = Rango de una muestra. La diferencia entre los valores más grande y más pequeño en la muestra.

s = Desviación estándar de una muestra aleatoria.

σ = Desviación estándar poblacional.

$\hat{\sigma}$ = Estimador de σ.

LSC = Límite superior de control para una gráfica de control.

$X = \begin{cases} \textit{Uso de la gráfica de control:} \text{ valor de una medición iindividual proveniente de la población.} \\ \textit{Uso del muestreo de aceptación:} \text{ número de artículos defectuosos observados en una muestra de } n \text{ artículos.} \end{cases}$

$\overline{X}$ = Promedio aritmético de una muestra aleatoria de n mediciones independientes.

Z = variable normal estándar.

$z_{\alpha/2}$ = Un número tal que la probabilidad de observar un valor de Z que sobrepasase a $z_{\alpha/2}$ es $\alpha/2$.

Bibliografía

Aguayo, R. *Dr. Deming: The American Who Taught the Japanese about Quality*. Nueva York: Lyle Stuart, 1990.

Alexander, F. "ISO 14001: What Does It Mean for IE's". *IIE Solutions* 2 (enero 1996), pp. 15-18.

Alexander, T. "Cray's Way of Staying Super-Duper". *Fortune,* marzo 18, 1985, p. 76.

Baker, K. R. "Two Process Models in the Economic Design of an *X* Chart". *AIIE Transactions* I3 (1971), pp. 257-263.

Boothroyd, G. y P. Dewhurst. *Product Design for Assembly.* Wakefield, RI: Boothroyd Dewhurst, Inc., 1989.

Boothroyd, G., P. Dewhurst y W. A. Knight. *Product Design for Manufacturing.* Nueva York: Marcel Dekker, 1994.

Business Week. "The Quality Imperative". Número especial dedicado a la calidad. Nueva York: McGraw-Hill, 1991.

Cohen, L. *Quality Function Deployment: How to Make QFD Work for you.* Reading, MA: Addison Wesley, 1995.

Crosby, P. B. *Quality Is Free.* Nueva York: McGraw-Hill, 1979.

Crosby, P. B. *Quality without Tears: The Art of Hassle-Free Management.* Nueva York: McGraw-Hill, 1984.

Dahan, E. "Note on Listening to the Customer, Part I". Nota del profesor, Graduate School of Business, Stanford University, Stanford, CA, 1995.

DeGroot, M. H. *Probability and Statistics.* 2a. ed. Reading, MA: Addison Wesley, 1986.

Dehnad, K. *Quality Control, Robust Design, and the Taguchi Method.* Pacific Grove, CA: Wadsworth Cole, 1989.

Duncan, A. J. "The Economic Design of *X* Charts Used to Maintain Current Control of a Process". *Journal of the American Statistical Association* 51 (1956), pp. 228-242.

Duncan, A. J. *Quality Control and Industrial Statistics.* 5a. ed. Nueva York: McGraw-Hill/Irwin, 1986.

Dvorak, P. "Manufacturing Puts a New Spin on Design". *Machine Design* 67 (agosto 22, 1994), pp. 67-74.

Feigenbaum, A. V. *Total Quality Control.* 3a. ed. Nueva York: McGraw-Hill, 1983.

Fortune. "How Jack Welch Keeps the Ideas Coming at GE". Agosto 13, 1991.

Garvin, D. A. "Quality on the Line". *Harvard Business Review* 61 (1983), pp. 64-75.

Garvin, D. A. *Managing Quality.* Nueva York: Free Press, 1988.

Grant, E. L. y R. S. Leavenworth. *Statistical Quality Control.* 6a. ed. Nueva York: McGraw-Hill, 1988.

Green, P. y V. R. Rao. "Conjoint Measurement for Quantifying judgmental Data". *Journal of Marketing Research* 8 (1971), pp. 355-363.

Groocock, J. M. *The Chain of Quality: Market Dominance through Product Superiority.* Nueva York: John Wiley & Sons, 1986.

Herron, D. A. Comunicación privada, 1985.

Hillkirk, J. "Europe Upstages Quest for Baldrige Award". *USA Today,* abril 22, 1992.

Ishikawa, K. *What Is Total Quality Control? The Japanese Way.* Traducido del japonés por D. J. Lu. Englewood Cliffs, N.J: Prentice Hall, 1985.

Kackar, R. N. "Off-Line Quality Control, Parameter Design, and the Taguchi Method". *Journal of Quality Technology* 17 (1985), pp. 176-188.

Kolesar, P. "Scientific Quality Management and Management Science". *Handbooks in Operations Research and Management Science,* vol. 4, *Logistics* of *Production and Inventory,* S. Graves, A. H. G. Rinnooy Kan; y P. Zipkin, comps. Capítulo 13. Amsterdam: North Holland, 1993.

Kumar, S. e Y. Gupta. "Statistical Process Control at Motorola's Austin Assembly Plant". *Interfaces* 23, núm. 2 (marzo-abril 1993), pp. 84-92.

Kuster, T. "ISO 9000: A 500-lb Gorilla?" *Metal Center News* 35, núm. 10 (septiembre 1995), pp. 5-6.

Lawler, E. E. y S. A. Mohrman. "Quality Circles after the Fad". *Harvard Business Review* 63 (1985), pp. 65-71.

Leonard, H. y E. Sasser. "The Incline of Quality". *Harvard Business Review* 60 (1982), pp. 163-171.

Logothetis, N. y H. P. Wynn. *Quality through Design.* Oxford: Clarendon Press, 1989.

Martin, J. "Ignore Your Customer". *Fortune,* mayo 1, 1995, pp. 121-126.

Miller, J. G., A. D. Meyer y J. Nakane. *Benchmarking Global Manufacturing.* Nueva York: McGraw-Hill/Irwin, 1992.

Montgomery, D. C. *Introduction to Statistical Quality Control.* Nueva York: John Wiley & Sons, 1985.

Motorola, Inc. "Motorola Corporate Quality System Review Guidelines". Edición de marzo de 1991. Citado en S. Kumar e Y. Gupta, "Statistical Process Control at Motorola's Austin Assembly Plant". *Interfaces* 23, núm. 2 (marzo-abril 1993), pp. 84-92.

Pierce, R. J. *Leadership Perspective, and Restructuring for Total Quality.* Milwaukee: ASQC Quality Press, 1991.

Pugh, S. *Total Design.* Workingham, Inglaterra: Addison Wesley, 1991.

Ross, S. M. *Applied Probability Models with Optimization Applications.* San Francisco: Holden Day, 1970.

Schoeffler, S., R. D. Buzzell y D. F. Heang. "Impact of Strategic Planning on Profit Performance". *Harvard Business Review* 52 (marzo-abril 1974).

Shewhart, W. A. *Economic Control of the Quality of manufactured Product.* New York: D. Van Nostrand, 1931.

Taguchi, G., A. E. Elsayed y T. Hsiang. *Quality Engineering in Production Systems.* Nueva York: McGraw-Hill, 1989.

Transportation and Distribution 36, núm. 5 (mayo 1995), pp. 26-28.

Ulrich, K. T. y S. D. Eppinger. *Product Design and Development.* Nueva York: McGraw-Hill, 1995.

Velury, J. "Integrating ISO 9000 into the Big Picture". *IIE Solutions* 1 (octubre 1995), pp. 26-29.

Capítulo **Doce**

Confiabilidad y facilidad de mantenimiento

Panorama general del capítulo

Propósito

Apreciar la importancia de la confiabilidad, comprender los mecanismos por los que fracasan los productos y entender las matemáticas que subyacen a estos procesos.

Puntos clave

1. *Preparación.* Los temas de este capítulo (teoría de la confiabilidad, garantías y remplazo por edad) se tratan en muy pocas ocasiones en los textos sobre operaciones. Se incluyen aquí debido a su importancia y relevancia para el movimiento de la calidad. Sin embargo, las matemáticas de la confiabilidad son complejas. Es necesario contar con un entendimiento básico de las variables aleatorias, densidad de probabilidad y funciones de distribución, además de procesos estocásticos elementales. Varios de estos métodos se utilizaron también en el capítulo 5 y en el Suplemento 2 sobre teoría de colas, que aparece en el capítulo 8. Sugiero al lector que revise cuidadosamente la discusión sobre distribución exponencial que ahí se presenta.
2. *Confiabilidad de un solo componente.* Considere un solo artículo cuyo tiempo de falla no puede predecirse con anticipación; esto sería una variable aleatoria, T. Suponemos que conocemos tanto la función de distribución como las funciones de densidad de T: $F(t)$ y $f(t)$, respectivamente. Varias cantidades importantes asociadas con T incluyen la función de supervivencia $R(t) = 1 - F(t)$, que es la probabilidad de que el artículo sobreviva más allá de t, y la función de tasa de falla, que se define como $r(t) = f(t)/R(t)$.

   Un caso importante se presenta cuando la función de tasa de falla es una constante independiente de t. Esto da como resultado la distribución de tiempo de falla con una distribución exponencial. Ésta es la única que posee la propiedad de amnesia. En este contexto significa que el artículo no está mejorando ni empeorando con la edad. Las funciones de tasa de falla crecientes y decrecientes, respectivamente, representan los casos en los que la confiabilidad de un artículo está mejorando o empeorando con la edad. La distribución Weinbull es una opción usual para representar las funciones crecientes y decrecientes de tasa de falla.
3. *El proceso Poisson en el modelado de confiabilidad.* El proceso Poisson es quizás el proceso estocástico más importante para aplicaciones. Cuando los tiempos entre fallas son independientes y las variables aleatorias exponenciales están distribuidas de manera idéntica (IID), se puede inferir que el número total de falla hasta cualquier punto t sigue una distribución Poisson y las n fallas siguen una distribución Earlang. Debido a que la

distribución exponencial es amnésica, el proceso describe con exactitud los eventos que ocurren completamente al azar a lo largo del tiempo.

4. *Confiabilidad de equipo complejo.* Los artículos propensos a fallar generalmente están construidos por más de un componente. En un sistema en serie, el sistema falla cuando cualquiera de los componentes falla. Una tercera posibilidad es un sistema K fuera de N. Aquí el sistema funciona siempre que lo haga cuando menos un componente K. En esta sección mostramos cómo obtener las distribuciones del tiempo de falla para esos sistemas, con base en las distribuciones del tiempo de falla para los componentes que constituyen los sistemas.
5. *Modelos de mantenimiento.* Por mantenimiento preventivo se entiende el remplazo de un artículo antes de que falle. Claramente, esto sólo tiene sentido para los artículos que tienen mayores posibilidades de fallar con la edad. Al remplazar los artículos con regularidad antes de que fallen se pueden evitar las interrupciones que resultan de las fallas no planeadas. De acuerdo con lo que se conoce sobre los mecanismos de falla de los artículos y los costos de remplazos planeados y no planeados, se pueden implementar estrategias óptimas de remplazo. El caso más sencillo da una fórmula para los tiempos óptimos de remplazo, que es muy similar a la fórmula EOQ que aparece en el capítulo 4.
6. *Garantías.* Una garantía es un convenio entre el comprador y el vendedor de un artículo en el que el vendedor está de acuerdo en proporcionar la restitución al comprador en el caso de que el artículo fallara dentro del periodo de garantía. Las garantías son comunes para casi todos los artículos de consumo y las garantías extendidas constituyen un excelente negocio. En esta sección, examinamos dos clases de garantías: la garantía de remplazo gratuito y la garantía de remplazo prorrateado. En la primera, como su nombre lo indica, el vendedor está de acuerdo en remplazar el artículo cuando falla durante el periodo de garantía. En el caso de la garantía prorrateada, el monto de restitución depende del tiempo que resta de garantía (las garantías prorrateadas son comunes para las llantas, por ejemplo, donde la devolución depende de la rodadura restante de la llanta).
7. *Confiabilidad del software.* El software está jugando un papel cada vez más importante en nuestras vidas. Con el explosivo crecimiento de las computadoras personales, el mercado del software para este tipo de dispositivos es enorme. Microsoft aprovechó este crecimiento para convertirse en una de las compañías más grandes a nivel mundial antes de cumplir diez años de su fundación. Sin embargo, la industria del software siempre es mucho más grande que las computadoras personales. Grandes bases de datos, como las manejadas por las autoridades fiscales o el Departamento de Tránsito de su estado, requieren de sistemas masivos para recuperar información. Algunos predijeron que el sistema de defensa de la Guerra de las Galaxias de Ronald Reagan estaba condenado al fracaso porque era imposible diseñar un software confiable para él. Las fallas del software pueden ser tan catastróficas como las del hardware, haciendo fracasar a los principales sistemas.

Nuestros sistemas críticos continúan creciendo y volviéndose cada vez más complejos. Conforme aumenta la complejidad, la confiabilidad se ve amenazada. Durante la semana del 11 de agosto de 2003, una falla en el suministro eléctrico cerca de Cleveland, Ohio, provocó uno de los peores apagones en la historia estadounidense. Prácticamente toda la costa oriental de Estados Unidos y partes de Canadá se vieron afectadas y fueron varios los fallecimientos de personas atribuidos al problema. ¿Cómo pudo haber ocurrido esto? La respuesta es que la red eléctrica estadounidense está enlazada a lo largo de todo el país y puede caerse incluso por problemas muy sencillos. Dichos sistemas deben diseñarse con mayor atención a su confiabilidad. Conforme crece la población y los sistemas básicos se vuelven cada vez más complejos, mayores son las probabilidades de enfrentar catástrofes de este tipo. Todos los días dependemos de la confiabilidad de nuestra infraestructura.

En los últimos años, la calidad ha sido un aspecto clave en la administración de operaciones. El gran éxito de los japoneses se ha atribuido en gran parte a la calidad de sus artículos manufacturados. La calidad tiene varias dimensiones (véase el capítulo 11, sección 13), y ciertamente la confiabilidad es un componente clave. En la tabla 11-7 (capítulo 11) se ven las prioridades en la competencia para Europa, Japón y Estados Unidos, durante los cinco años siguientes. La confiabilidad en el producto fue clasificada como número uno en el grupo de empresas japonesas estudiadas.

Cuando pensamos en el éxito económico de Japón, lo primero que nos viene a la mente es la industria automotriz. Los fabricantes japoneses han tenido una parte cada vez mayor en el mercado de Estados Unidos durante las últimas dos décadas, aun durante periodos con cuotas de importación o de tasas de cambio desfavorables. ¿Por qué han tenido tanto éxito los japoneses en Estados Unidos? Es probable que la razón fundamental sea la calidad percibida en el producto, debido a lo cual hay tantos estadounidenses que optan por comprar automóviles japoneses. Pero, ¿qué dimensión de la calidad es la más importante? Una respuesta probable es la confiabilidad del producto. Las encuestas anuales que efectúa la Unión del Consumidor atestiguan la confiabilidad excepcional de los automóviles fabricados por japoneses.

La confiabilidad, como un campo individual, se separó del control estadístico de la calidad en la década de 1950, con el crecimiento de las industrias aeroespacial y electrónica en Estados Unidos. El Departamento de Defensa se interesó mucho en estudios de confiabilidad al apreciar, dolorosamente, que había un grave problema con la fiabilidad de componentes y sistemas militares. Garvin (1988) informa que en 1950 sólo la tercera parte de los aparatos electrónicos de la Marina trabajaban bien en determinado momento, y que por cada tubo al vacío en posibilidad de funcionar, había nueve en los almacenes o en los pedidos. Según Amstadter (1971), el costo anual de mantener algunos sistemas militares en condición operativa ha costado hasta 10 veces el valor original del equipo.

¿Cuál es la diferencia entre el control estadístico de la calidad y la confiabilidad? El control estadístico de la calidad se ocupa de monitorear los procesos para asegurar que el producto manufacturado se apegue a las especificaciones. Las variables aleatorias de interés son la cantidad de defectos y el grado de variación de apego a las normas. En la confiabilidad se examina el desempeño de un producto a través del *tiempo.* Las variables aleatorias de interés son la cantidad de tiempo transcurrido entre fallas, una vez que el producto se pone en servicio. O'Connor (1985) ha sugerido una definición de confiabilidad, que subraya su relación estrecha con la calidad: la confiabilidad es un concepto de la calidad basado en el tiempo. La confiabilidad también es la probabilidad de que un producto funcione en forma correcta durante un periodo dado, en su desempeño previsto (Amstadter, 1971).

Un revisor de la primera edición en inglés de este libro dijo que, desde el punto de vista del alumno, la confiabilidad es "un asunto muy apegado a la ingeniería, que casi no figura en los titulares de *The Wall Street Journal"*. Esto no puede estar más lejos de la verdad. Tres de los desastres más importantes de los tiempos recientes fueron resultado de fallas de confiabilidad: los accidentes de las plantas nucleares de Three Mile Island en Pennsylvania, Estados Unidos, y en Chernobyl, en la antigua Unión Soviética, así como la dramática pérdida del *Challenger,* el peor desastre del programa espacial estadounidense. La confiabilidad nos concierne a todos diariamente. Nuestras vidas dependen de la confiabilidad de los automóviles, los transportes colectivos y los aviones. Nuestra calidad de vida depende de la confiabilidad de que se genere electricidad, de los teléfonos y de las computadoras. Nuestra salud depende de la confiabilidad de los sistemas de control de contaminación, los de calefacción y acondicionamiento de aire, y los de atención médica de urgencia.

La confiabilidad y el riesgo se relacionan en forma estrecha. Los riesgos debidos a baja confiabilidad conciernen tanto al productor como al consumidor. Algunos aspectos del riesgo, desde el punto de vista del productor, comprenden:

1. *La competencia.* La confiabilidad del producto es un componente importante de la calidad que percibe el consumidor. Los productos muy poco confiables no merecen la lealtad del cliente y, finalmente, desaparecen.

2. *Los requerimientos del consumidor.* El gobierno estadounidense requería sistemas de armamento con niveles claramente especificados de confiabilidad, cuando advirtió que los costos de mantenimiento de esos sistemas se estaban haciendo prohibitivos. Hoy son comunes los requisitos establecidos por el comprador.
3. *Los costos de garantía* y *servicio.* Las garantías, que se describirán en la segunda parte de este capítulo, son una carga financiera apreciable para el fabricante, cuando sus productos no son confiables. Las empresas automotrices estadounidenses otorgan garantías prolongadas, desde hace varios años, como incentivo al consumidor para poder aumentar las ventas (la más notable fue la de siete años o 70 000 millas de Chrysler). No hay pruebas de que esas garantías se hayan acompañado de más confianza. Como resultado de ello, esos programas fueron muy costosos para las empresas automotrices.
4. *Los costos de responsabilidad civil.* En gran parte como consecuencia de los esfuerzos de Ralph Nader, el Congreso de Estados Unidos ha promulgado leyes que hacen que el fabricante sea responsable de las consecuencias de las fallas en el desempeño del producto debidas a mal diseño o mala manufactura. Las pérdidas por responsabilidad civil han provocado que los costos de la mala confiabilidad se hayan desplazado del consumidor al productor.

Algunos de los riesgos de la mala confiabilidad que asume el consumidor son:

1. *Seguridad.* No hay duda que las fallas en el equipo pueden ocasionar la muerte. Cada año mueren aproximadamente 50 000 estadounidenses por accidentes en las carreteras de ese país. Sin duda, una parte de estos percances puede atribuirse a fallas mecánicas. Los accidentes aéreos provocan la muerte, y muchos de ellos son resultado de fallas en el equipo. Las fallas de las plantas nucleares de Three Mile Island y de Chernobyl causaron muertes y daños a los humanos por exposición a la radiación. La confiabilidad y la seguridad están estrechamente relacionadas.
2. *Incomodidad.* Aun cuando muchas fallas no traen como consecuencia la muerte, pueden provocar frustración y demoras. Son comunes las demoras en los aeropuertos debidas a que algún equipo a bordo del avión no funciona bien. Las descomposturas automotrices pueden dejar varados a los conductores durante horas. Las fallas en el equipo de comunicación, de cómputo o de generación de electricidad pueden dejar paradas a las empresas.
3. *Costo.* En última instancia, la baja confiabilidad les cuesta a todos. Por esta razón, los consumidores están de acuerdo en pagar algo adicional en productos más fiables. Los japoneses lo han aprendido bien, y todo hace suponer que continuarán con la estrategia de aumentar su participación en el mercado produciendo artículos más confiables que los de sus competidores.

¿Por qué estudiar confiabilidad? Necesitamos comprender las leyes de probabilidad que rigen a los patrones de falla para diseñar mejores procesos que produzcan sistemas confiables. Un análisis incorrecto puede tener consecuencias desastrosas. Por ejemplo, en el llamado Informe Rasmussen (U. S. Nuclear Regulatory Commission, *Reactor Safety Study,* 1975) se predijo que pasarían siglos para que fuera probable un gran accidente de planta nuclear. Como desde entonces han ocurrido dos grandes accidentes, es claro que el análisis en este informe no es válido.

La confiabilidad es un asunto que concierne a la administración de operaciones desde dos puntos de vista. Primero, para implementar la administración de la calidad total debemos comprender cómo y cuándo fallan los productos. La confiabilidad continuará siendo un componente clave de la calidad. Para diseñar un sistema efectivo de entrega de calidad se requerirá comprender la aleatoriedad del comportamiento de las fallas. En segundo lugar, será necesario comprender las propiedades de las fallas del equipo que se usa en el proceso de manufactura. De esta manera podremos desarrollar políticas efectivas de mantenimiento para esos equipos.

12.1 CONFIABILIDAD DE UN SOLO COMPONENTE

Introducción a los conceptos de confiabilidad

Para que los lectores comprendan mejor y aprecien más algunas de las definiciones y conceptos que presentamos en este capítulo, esta sección comienza con un ejemplo.

Ejemplo *12.1* En 1970, el ejército estadounidense compró 1 000 capacitores eléctricos para usarlos en radiotransmisores de corto alcance. El ejército mantuvo registros detallados del patrón de fallas de esos capacitores, con los siguientes resultados:

Número de años en operación	1	2	3	4	5	6	7	8	9	10	>10
Número de fallas	220	158	121	96	80	68	47	40	35	25	110

Con base en estos datos, se desea estimar la distribución de probabilidad asociada con la falla de un capacitor seleccionado al azar.

Definiremos a la variable aleatoria T como el tiempo que el capacitor funciona hasta que falla. Podemos estimar la función de distribución acumulada de T, $F(t)$, con los datos citados. Usando símbolos, $F(t) = P\{T \leq t\}$, y con palabras, $F(t)$ es la probabilidad de que un componente elegido al azar falle en o antes del tiempo t.

Para estimar a $F(t)$ a partir de los datos, se determinará la cantidad acumulada de fallas y la proporción del total que representa esa cantidad cada año. Entonces

Años de servicio	1	2	3	4	5	6	7	8	9	10	>10
Fallas acumuladas	220	378	499	595	675	743	790	830	865	890	1 000
Proporción del total	.220	.378	.499	.595	.675	.743	.790	.830	.865	.890	1.0

Las proporciones son estimados de $F(t)$ para $t = 1, 2, \ldots$ Esas probabilidades se pueden usar en forma directa para calcular diversas cantidades de interés, considerando que T es una variable aleatoria discreta, o bien se pueden usar para estimar los parámetros de una distribución continua. Usaremos la versión discreta para contestar varias preguntas sobre la vida de los capacitores. Por ejemplo, supongamos que se intenta determinar:

a) La probabilidad de que un capacitor tomado al azar dure más de 5 días.

b) La proporción de los 1000 capacitores originales que trabajaron y que fallaron el año 6.

c) La proporción de los capacitores que sobreviven cuando menos 5 años y fallan en el año 6.

d) La proporción de los capacitores que sobreviven cuando menos 8 años y fallan en el año 9.

Solución

a) $P\{T > 5\} = 1 - P\{T \leq 5\} = 1 - .675 = .325.$

b) $P\{T = 6\} = P\{T \leq 6\} - P\{T \leq 5\} = .743 - .675 = .068$

c) A primera vista, parece que esta pregunta es igual a la del inciso *b*). Sin embargo, hay una diferencia importante. En el inciso (*b*) se pide la proporción del conjunto original de capacitores que falla en el año 6, mientras que en el inciso (*c*) se pide la proporción de capacitores que duran cinco años, cuando menos, y que fallan en el año 6. Se trata de una probabilidad *condicional*.

$$P\{T = 6 \mid T > 5\} = \frac{P\{T = 6, T > 5\}}{P\{T > 5\}} = \frac{P\{T = 6\}}{P\{T > 5\}} = \frac{.068}{.325} = 0.209.$$

Observe que los eventos $\{T = 6, T > 5\}$ y $\{T = 6\}$ son equivalentes, porque $\{T = 6\} \subset \{T > 5\}$.

d) $$P\{T = 9 \mid T > 8\} = \frac{P\{T = 9\}}{P\{T > 8\}} = \frac{.035}{.170} = 0.206,$$

que casi es exactamente igual a la proporción de capacitores que sobreviven más de 5 años y fallan en el año 6. De hecho, la proporción de componentes que sobreviven n años y que fallan en el año $n + 1$ se aproxima mucho a .20, para $n = 0, 1, 2, \ldots, 9$. Como veremos después, esto da como resultado una distribución de fallas con ciertas propiedades excepcionales.

Notación y definiciones preliminares

En el ejemplo 12.1 se introdujeron muchos conceptos nuevos. En esta sección se formalizarán las definiciones, para poder usarlas en este capítulo.

Como antes, definimos la variable aleatoria T como la vida del componente. Se supone que T tiene una función $F(t)$ de distribución acumulada dada por

$$F(t) = P\{T \leq t\}.$$

En lo que sigue consideraremos que $F(t)$ es una función diferenciable de t, de modo que exista la función $f(t)$ de densidad de probabilidades expresada por la ecuación

$$f(t) = \frac{dF(t)}{dt}.$$

Además de las funciones de distribución y de densidad de la variable aleatoria T, nos interesan otras funciones afines. Una es la función de confiabilidad (que también se llama función de supervivencia). La función de confiabilidad del componente, se representa por $R(t)$ y está dada por:

$$R(t) = P\{T > t\} = 1 - F(t).$$

En palabras, $R(t)$ es la probabilidad de que un componente nuevo sobreviva más allá del tiempo t. Observe que esto implica que $F(t)$ es la probabilidad de que un componente nuevo *no* sobreviva más allá del tiempo t.

Considere la siguiente probabilidad condicional:

$$P\{t < T \leq t + s \mid T > t\}.$$

Ésta es la probabilidad condicional de que un componente nuevo falle entre t y $t + s$, dado que dura más de t. Podemos imaginar que esta probabilidad condicional es lo siguiente: t es ahora, y s es un incremento del tiempo hacia el futuro. El evento $\{T > t\}$ indica que el componente ha sobrevivido hasta el presente o, en otras palabras, que todavía funciona. El evento condicional $\{t < T \leq t + s \mid T > t\}$ indica que el componente funciona ahora, pero que fallará antes de que pasen otras s unidades de tiempo.

Recuerde, de la teoría elemental de probabilidad, que para cualesquiera eventos A y B,

$$P\{A \mid B\} = \frac{P\{A \cap B\}}{P\{B\}}.$$

En el caso especial en que $A \subset B$, $A \cap B = A$, de modo que

$$P\{A \mid B\} = \frac{P\{A\}}{P\{B\}} \qquad \text{cuando } A \subset B.$$

Se identifican los eventos como sigue: $A = \{t < T \leq t + s\}$ y $B = \{T > t\}$. Con algo de razonamiento se llega a que en este caso particular, $A \subset B$, de modo que

$$P\{t < T \leq t + s \mid T > t\} = \frac{P\{t < T \leq t + s\}}{P\{T > t\}} = \frac{F(t + s) - F(t)}{R(t)}.$$

Dividiremos entre s y haremos que s tienda a cero.

$$\lim_{s \to 0} \frac{1}{-} \frac{F(t+s) - F(t)}{R(t)} = \frac{f(t)}{R(t)}.$$

Esta relación es una cantidad fundamental en la teoría de la confiabilidad.

Definiremos a

$$r(t) = \frac{f(t)}{R(t)}.$$

Llamaremos a $r(t)$ la *función de tasa de fallas.* Su deducción es la mejor manera de comprender lo que significa la función de tasa de fallas. Cuando s es positiva, la probabilidad condicional que se usó para llegar a $r(t)$ es la probabilidad de que un componente haya sobrevivido hasta el momento t, y que falle entre los tiempos t y $t + s$. Dividir entre s y dejar que s tienda a 0 es lo mismo que determinar una primera derivada. Por consiguiente, la función de tasa de fallas es la rapidez con que cambia la probabilidad condicional de falla en el momento t. Se puede considerar como una medida de la probabilidad de que un componente que ha sobrevivido hasta el tiempo t falle en el siguiente instante.

La función de tasa de fallas es una cantidad fundamental en la teoría de la confiabilidad, pero, al igual que la función de densidad de probabilidades, no tiene interpretación física directa. Sin embargo, para valores suficientemente pequeños de Δt, el término $r(t)\Delta t$ es la probabilidad de que un artículo que sobrevive hasta el tiempo t falle entre t y $t + \Delta t$. ¿Cómo se determina si Δt es suficientemente pequeño? En general, Δt debe ser pequeño en relación con la vida de un componente característico. Sin embargo, la única forma de tener la seguridad de que $r(t)\Delta t$ sea una buena aproximación a esta probabilidad condicional es calcular directamente $P\{t < T \leq t + \Delta t | T > t\}$.

Ejemplo ***12.2***

El intervalo de tiempo que determinada pieza de un equipo funciona hasta que falla es una variable aleatoria cuya función de distribución acumulada es

$$F(t) = 1 - e - 0.043t^{2.6}.$$

Considere lo siguiente:

a) La función de tasa de fallas.

b) La probabilidad de que el equipo trabaje durante más de 5 años sin fallar.

c) Suponga que se ponen en servicio 100 piezas en el año 0. ¿Qué fracción de unidades que sobreviven 4 años falla en el año 5? ¿Se puede estimar con exactitud esa proporción sólo con una función de tasa de fallas?

d) ¿Qué fracción de unidades que sobreviven 4 años fallan en el primer mes del año 5? ¿Se puede estimar eso con exactitud con la función de tasa de falla?

Solución

a)

$$f(t) = \frac{dF(t)}{dt} = -e^{-0.043t^{2.6}} \frac{d}{dt}(-0.043t^{2.6})$$

$$= (0.043)(2.6)t^{1.6} e^{-0.043t^{2.6}}$$

$$= 0.1118t^{1.6} e^{-0.043t^{2.6}}$$

Como $R(t) = 1 - F(t) = e^{-0.043t^{2.6}}$, entonces

$$r(t) = \frac{f(t)}{R(t)} = 0.1118t^{1.6}.$$

b) $$P\{T > 5\} = R(5) = e^{-0.043(5)^{2.6}} = e^{-2.8235} = 0.0594.$$

c) La calcularemos en forma directa y compararemos el resultado con $r(t)\Delta t$. La proporción de unidades que sobrevive 4 años y falla el año 5 es

$$P\{4 < T \leq 5 \mid T > 4\} = \frac{F(5) - F(4)}{R(4)} = \frac{R(4) - R(5)}{R(4)}$$

$$= \frac{e^{-0.043(4)^{2.6}} - e^{-0.043(5)^{2.6}}}{e^{-0.043(4)^{2.6}}}$$

$$= \frac{0.2059 - 0.0594}{0.2059}$$

$$= 0.7115.$$

Esto significa que, alrededor de 71% de las partes que sobreviven 4 años fallarán en el quinto año. Como 94% de las unidades fallan en los primeros 5 años de funcionamiento [$F(5) = .9406$], es probable que un valor de $\Delta t = 1$ año sea demasiado grande para que $r(t)\Delta t$ dé una buena aproximación. De hecho, se ve que $r(4)(1) = (0.1118)(4)^{1.6} = 1.0274$.

d) Como un mes corresponde a $\frac{1}{12} = 0.0833$ año, se debe calcular

$$P\{4 < T \leq 4.0833 \mid T > 4\} = \frac{F(4.0833) - F(4)}{R(4)} = \frac{0.2059 - 0.1887}{0.2059} = 0.0836.$$

Aquí, $\Delta t = \frac{1}{12}$ debe ser lo suficientemente pequeño como para poder usar la función de tasa de falla para estimar esta probabilidad. El resultado es $r(4)\,\frac{1}{12} = 0.0856$.

La ley exponencial de fallas

La distribución exponencial juega un papel fundamental en la teoría y la práctica de la confiabilidad, porque describe con exactitud las características de falla de muchos equipos en funcionamiento. Se puede deducir la ley exponencial en varias formas. Describiremos una deducción en la que se usa la función de tasa de fallas.

Sabemos que la función de tasa de fallas *r(t)* se expresa con la fórmula

$$r(t) = f(t)/R(t)$$

y es una medida de la posibilidad de que una unidad que haya estado funcionando t unidades de tiempo falle en el próximo instante. Examinemos el caso siguiente: $r(t) = \lambda$ para alguna $\lambda > 0$ constante. Esto indica que la probabilidad de que una unidad que está funcionando falle en el próximo instante es independiente de cuánto tiempo ha estado trabajando. Esto implica que la unidad no presenta síntomas de envejecimiento. Es igualmente probable que falle en el siguiente instante, cuando está nueva o cuando no lo está. Deduciremos la distribución de probabilidades de la vida T, que corresponde a una función de tasa constante de falla.

Podemos determinar una solución de la ecuación $r(t) = \lambda$ si notamos que $R(t) = 1 - F(t)$,

$$f(t) = \frac{dR(t)}{} = -R'(t).$$

Por consiguiente, la ecuación $r(t) = \lambda$ se puede escribir en la forma

$$\frac{-R'(t)}{R(t)} = \lambda$$

o bien

$$R'(t) = -\lambda R(t).$$

Ésta es la ecuación lineal de primer orden más sencilla. Su solución es

$$R(t) = e^{-\lambda t}.$$

Por consiguiente, la función de distribución $F(t)$ se expresa como sigue:

$$F(t) = 1 - e^{-\lambda t}$$

y la función de densidad $f(t)$ es

$$f(t) = \lambda e^{-\lambda t}.$$

A esta ecuación se le llama distribución exponencial. Depende del único parámetro λ, que representa una tasa de ocurrencia. Si T tiene la distribución exponencial con parámetro λ, representa la vida de un componente que no envejece con el tiempo; esto es, un componente que ha sobrevivido hasta el tiempo t_1 tiene la misma probabilidad de fallar en el siguiente instante que uno que ha sobrevivido hasta el tiempo t_2 sean cuales sean t_1 y t_2. El tiempo esperado de falla es $1/\lambda$. La desviación estándar del tiempo hasta la falla también es $1/\lambda$. Las funciones de densidad y distribución exponenciales aparecen en la figura 12-1.

Ejemplo *12.3*

Como la distribución exponencial tiene una función de tasa de fallas constante, es probable que la ley de falla para el capacitor descrito al principio del capítulo sea exponencial. Esto se debe a que, según vimos, la proporción de los capacitores que han sobrevivido durante n años y fallan en el año $n + 1$ fue igual para $n = 0, 1, \ldots$, que era 20%, aproximadamente. Para poder estimar el valor de λ, se observa que la proporción que falla en el primer año también es 20%, y esto hace que $F(1) = 0.20 = 1 - e^{-\lambda}$. Al despejar λ se obtiene $e^{-\lambda} = 0.8$, es decir $\lambda = -\ln(0.8) = 0.223$.

Podemos comparar la cantidad esperada de capacitores que fallarían si la distribución real de la vida fuera exponencial en función de la cantidad real que falló, para ver si con la distribución exponencial se obtiene un ajuste razonable a los datos.

Número de años en operación	1	2	3	4	5	6	7	8	9	10	>10
Número de fallas	220	158	121	96	80	68	47	40	35	25	110
Número de fallas esperadas bajo la ley exponencial con $\lambda = 0.223$	200	160	128	102	82	66	52	42	34	27	107

La cantidad esperada de fallas para cada año se obtiene multiplicando la probabilidad de falla para un año determinado, suponiendo que la ley es exponencial, por la cantidad total de unidades. Por ejemplo, se ve que la cantidad esperada de fallas para el año 3 se obtiene multiplicando 1 000 por $F(3) - F(2) = e^{-2\lambda} - e^{-3\lambda} = .1280$. Se aprecia claramente que hay buena concordancia entre las cantidades reales de fallas y las esperadas, indicando que la distribución exponencial permite tener un buen ajuste con los datos históricos observados.[1]

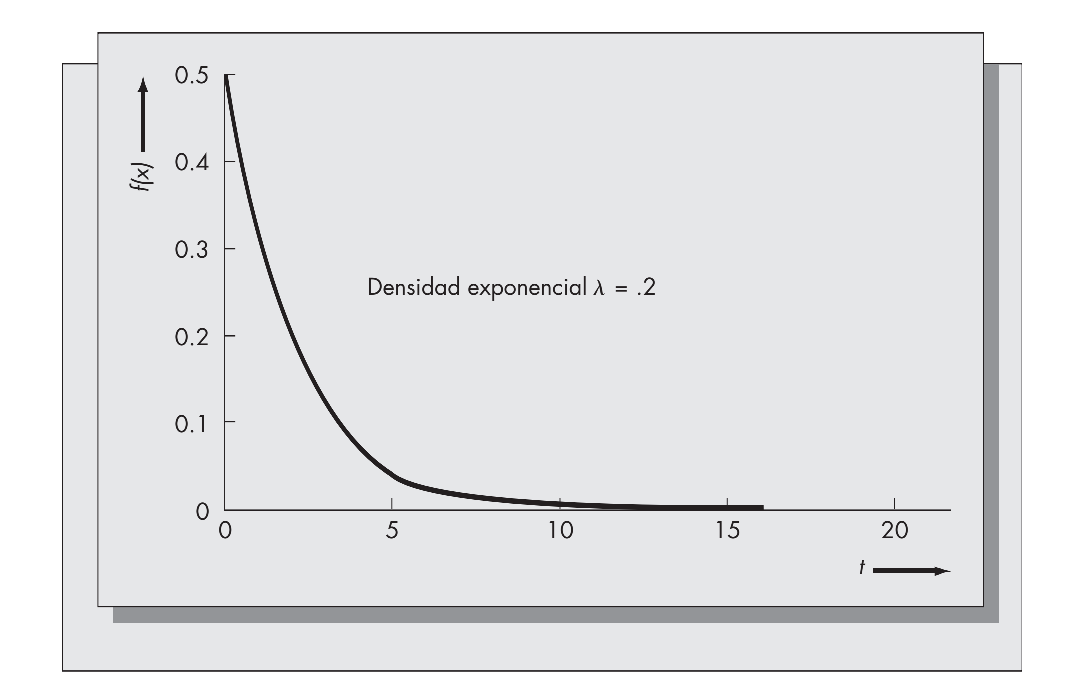
Ya podemos contestar varias preguntas acerca de los capacitores, si usamos directamente la distribución exponencial. Por ejemplo, supongamos que se desea calcular:

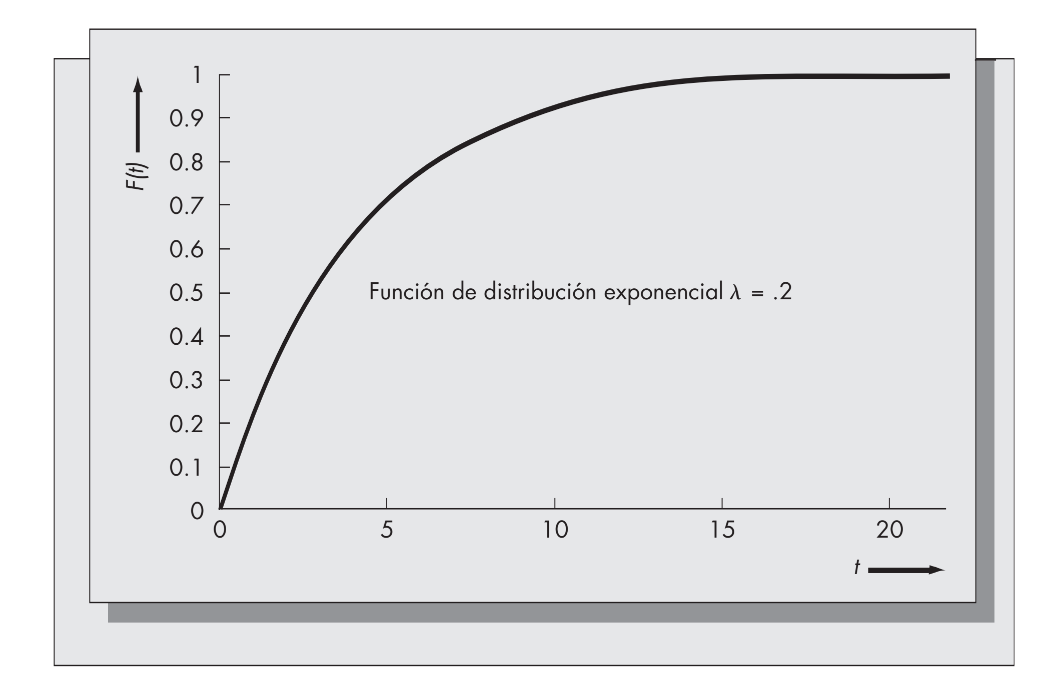
a) La probabilidad de que un capacitor tomado al azar dure más de 8 años.

b) La proporción de capacitores que sobreviven 3 años y que también sobreviven cuando menos otros 3 años más.

[1] Es fácil comprobar con una prueba de bondad de ajuste que estos datos se ajustan estrechamente a la distribución exponencial.

FIGURA 12-1
Funciones exponenciales de densidad y de distribución.





Solución

a) $P\{T > 8\} = e^{-\lambda t} = e^{-(.223)(8)} = e^{-1.784} = .1680.$

b) Se desea comparar $P\{T > 6 \mid T > 3\}$. Aplicando las leyes de la probabilidad condicional, se ve que

$$P\{T > 6 \mid T > 3\} = \frac{P\{T > 6,\, T > 3\}}{P\{T > 3\}}$$

$$= \frac{P\{T > 6\}}{P\{T > 3\}} = \frac{.2624}{.5122} = .5122.$$

No es una coincidencia que esto sea igual a la probabilidad no condicional de que un nuevo capacitor dure más de 3 años. Esta propiedad también se conoce como propiedad de amnesia de la distribución exponencial.

La propiedad de amnesia de la distribución exponencial se relaciona con la siguiente probabilidad condicional:

$$P\{T > t + s \,|\, T > t\}.$$

Ésta es la probabilidad de que el componente sobreviva más allá del tiempo $t + s$, dado que ha sobrevivido hasta el tiempo t. Si nos imaginamos que el tiempo o momento t es ahora mismo, entonces es la probabilidad de que un componente que está funcionando actualmente continúe funcionando cuando menos otras s unidades de tiempo. Si T sigue una ley exponencial de fallas, entonces

$$\begin{aligned} P\{T > t + s \,|\, T > t\} &= \frac{P\{T > t + s, T > t\}}{P\{T > t\}} \\ &= \frac{P\{T > t + s\}}{P\{T > t\}} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} \\ &= e^{-\lambda s} \\ &= P\{T > s\}. \end{aligned}$$

Observe que cuando $\{T > t + s\} \subset \{T > t\}$, los eventos $\{T > t + s, T > t\}$ y $\{T > t + s\}$ son equivalentes. La última ecuación, $P\{T > s\}$, es la probabilidad *no condicional* de que un nuevo componente dure cuando menos s unidades de tiempo. Esto es, hemos demostrado que si el componente ha estado trabajando durante t unidades de tiempo sin fallar, la probabilidad de que continúe funcionando cuando menos otras s unidades de tiempo es igual a la de que un componente nuevo funcione durante al menos s unidades de tiempo. Esto significa que no hay envejecimiento. La probabilidad de falla es independiente de cuánto tiempo ha estado trabajando el componente. Sin embargo, nuestro requisito fue que la distribución de los tiempos de vida fuera exponencial. De hecho, la distribución exponencial es la *única* distribución continua que posee la propiedad de amnesia; esto es, es la única para la que $P\{T > t + s \,|\, T > t\} = P\{T > s\}$.

Cuando decimos que un artículo falla en forma totalmente aleatoria estamos indicando que la ley de falla de ese artículo es exponencial. Los eventos que suceden totalmente al azar según pasa el tiempo se apegan a un *proceso de Poisson.* El proceso Poisson se analiza en la sección 12.3.

Problemas para la sección 12.1

1. Se pusieron en servicio, en forma simultánea, 300 tubos de rayos catódicos idénticos, el día 1 de enero de 1976, y fallaron como sigue, hasta el 31 de diciembre de 1988:

Año	1983	1984	1985	1986	1987	1988
Cantidad de fallas	13	19	16	34	21	38

Suponga que no hubo fallas antes de 1983.

a) Con base en estos datos, estime la función de distribución acumulada de un tubo de rayos catódicos tomado al azar.

Usando los resultados del inciso *a*), estime la probabilidad de que un tubo tomado al azar

b) Dure más de 5 años.

c) Dure más de 10 años.

d) Dure más de 12 años.

e) Habiendo sobrevivido 10 años, falle en el undécimo año de funcionamiento.

2. Suponga que la función de distribución acumulada de la vida de cierta parte es

$$F(t) = 1 - e^{-0.6t} - 0.6te^{-0.6t},$$

donde t se mide en años de operación continua.

a) Determine la función de confiabilidad.

b) Determine la función de tasa de fallas.

c) ¿Cuál es la probabilidad de que esta pieza falle en el primer año de operación?

d) ¿Cuál es la probabilidad de que esta pieza falle en el quinto año de operación?

e) ¿Qué proporción del equipo que sobrevive 4 años falla en el quinto año? (Calcule sin usar la función de tasa de fallas.)

f) ¿Se aproxima bien $r(4)$ a la respuesta del inciso (*e*)? ¿Por qué sí o por qué no?

g) ¿Qué proporción de las piezas que sobreviven 4 años falla en el primer mes del quinto año? (Calcule usando la función de tasa de fallas.)

3. Se pone en servicio una gran cantidad de piezas idénticas cuando el tiempo es 0. Esos artículos tienen la siguiente función de tasa de fallas:

$$r(t) = 1.105 + 0.30t,$$

en la que t se expresa en años de funcionamiento.

a) Deduzca $R(t)$ y $F(t)$.

b) Si en el momento $t = 1$ año siguen trabajando 300 piezas, ¿qué cantidad aproximada de piezas espera usted que falle entre el año 1 y el año 2?

c) ¿Produce el valor de $r(1)$, una buena aproximación a la probabilidad condicional calculada en el inciso (*b*)? ¿Por qué?

d) Repita el cálculo del inciso (*b*), pero determine la cantidad esperada de artículos que fallan desde $t = 1$ año hasta $t = 1$ año y una semana. ¿Se obtiene una aproximación razonable a la probabilidad condicional en este caso con $r(t)\Delta t$? ¿Por qué?

4. Un microprocesador que controla al sintonizador en las TV a color falla totalmente al azar, esto es, siguiendo la distribución exponencial. Suponga que la probabilidad de que un microprocesador que ha sobrevivido durante k años falle en el año $k + 1$ es igual a .0036. ¿Cuál es la función de distribución acumulada del tiempo hasta la falla, es decir, de la vida del microprocesador?

5. Un transductor de presión regula un sistema de control de clima en una fábrica. Este transductor falla apegándose a una distribución exponencial con la tasa de una falla cada 5 años, en promedio.

a) ¿Cuál es la función de distribución acumulada del tiempo hasta la falla?

b) ¿Cuál es la probabilidad de que un transductor tomado al azar funcione durante 8 años sin fallar?

c) ¿Cuál es la probabilidad de que un transductor que ha funcionado 8 años continúe trabajando otros 8 años sin fallar?

6. Para el transductor de presión del problema anterior, use la función de tasa de fallas para estimar la probabilidad de que un transductor que ha estado trabajando durante 6 años falle en el séptimo año. ¿Qué tanto se aproxima a la respuesta exacta?

12.2 TASAS DE FALLAS CRECIENTES Y DECRECIENTES

Aunque la función de tasa de falla constante, que conduce a la ley exponencial, es importante en la teoría de confiabilidad, también hay otras importantes leyes de falla. La mayoría de nosotros estamos más familiarizados con artículos que poseen funciones crecientes de tasa de falla. Esto es, resulta más probable que fallen a medida que envejecen. También se pueden hallar funciones decrecientes de tasa de falla. Con frecuencia, los productos nuevos tienen una alta frecuencia de fallas debido a la fase de "quemado", en la que salen los artículos defectuosos de la población.

Una clase importante de funciones de tasa de falla, que pueden ser crecientes o decrecientes, tiene la forma

$$r(t) = \alpha\beta t^{\beta-1} \qquad \text{donde } \alpha \text{ y } \beta > 0.$$

En este caso, $r(t)$ es una función polinomial en la variable t que depende de los dos parámetros α y β. Cuando $\beta > 1$, $r(t)$ es creciente, y cuando $0 < \beta < 1$, $r(t)$ es decreciente. En la figura 12-2 se presentan funciones típicas de tasa de falla para estos casos.

Esta forma de $r(t)$ dará como resultado otra ecuación diferencial en $R(t)$. Se puede demostrar que en este caso la solución es

$$R(t) = e^{-\alpha t^{\beta}} \qquad \text{para toda } t \geq 0$$

o bien

$$F(t) = 1 - e^{-\alpha t^{\beta}} \qquad \text{para toda } t \geq 0.$$

Esta distribución se conoce como distribución de Weibull. Depende de los dos parámetros α y β, y como ya se vio, cuando $0 < \beta < 1$, corresponde a la vida de un artículo con tasas decrecientes de falla, y cuando $\beta > 1$, corresponde al tiempo de vida de un artículo con tasas

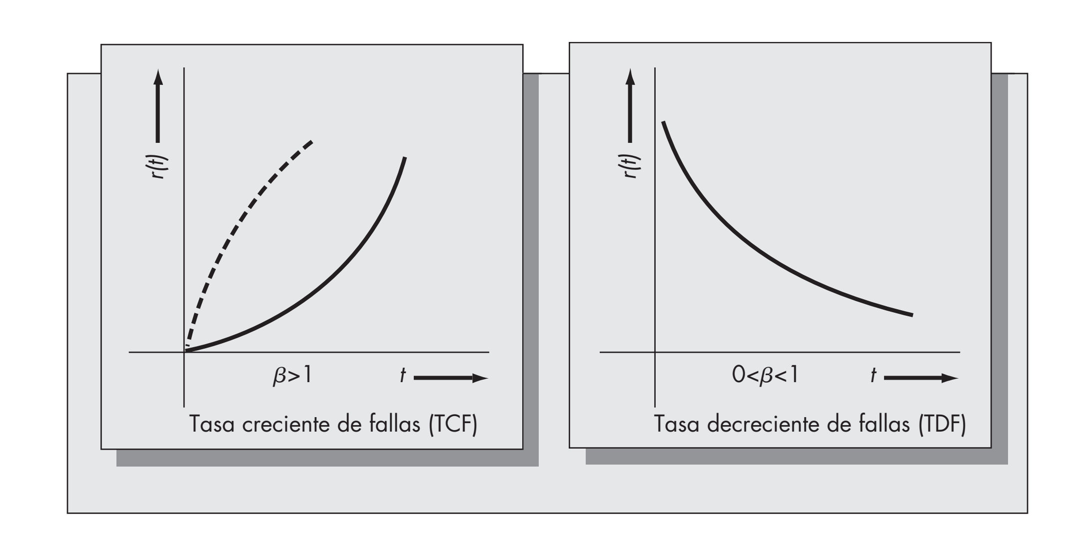


FIGURA 12-2
Funciones de tasa de fallas para la distribución de Weibull del tiempo de vida.

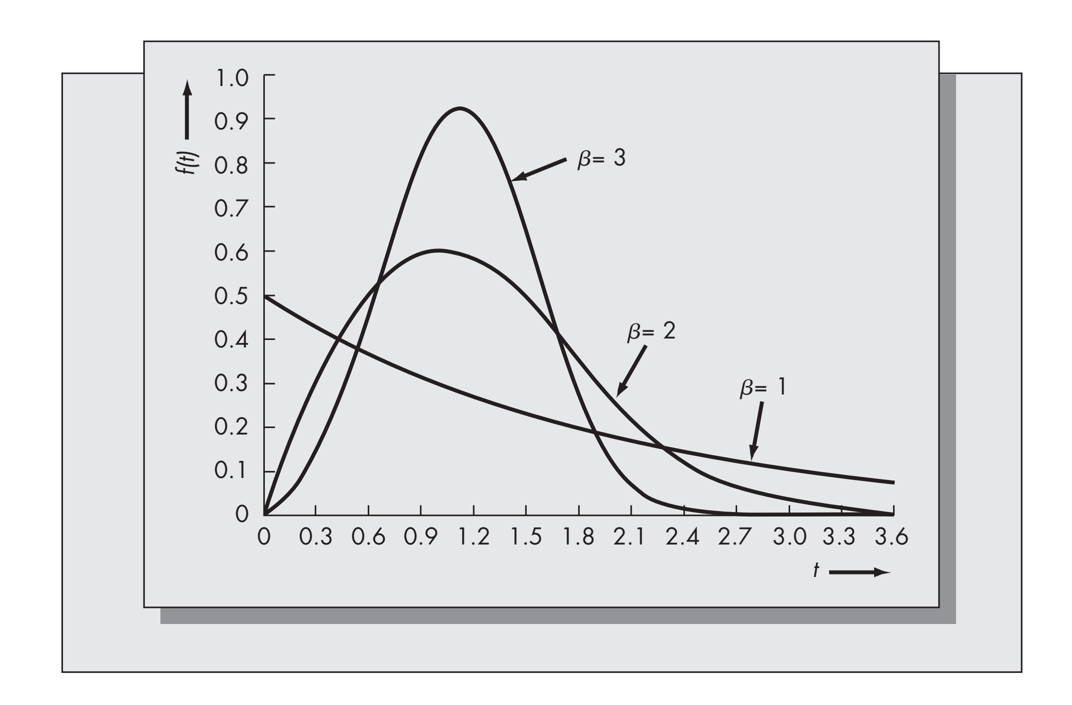


FIGURA 12-3
Densidades Weibull para diversos valores de β ($\alpha = 0.5$).

crecientes de falla. Como con frecuencia se cumple que las funciones empíricas de frecuencia de falla (es decir, las que se observan según datos de pruebas) se aproximan mucho por medio de polinomios, la distribución de Weibull es una descripción precisa de la ley de falla de muchos equipos. Obsérvese que cuando $\beta = 1$, la distribución de Weibull es igual a la exponencial. En la figura 12-3 aparecen varias densidades de Weibull.

Ejemplo *12.4*

Un fabricante de equipo de copiado incluye una garantía de reparación con cada copiadora. Casi todo su equipo presenta una tasa creciente de fallas. Con base en datos históricos, la tasa de fallas del modelo 25cc7 se describe fielmente con la función $r(t) = 2.7786t^{1.3}$, donde t se expresa en meses de trabajo continuo. ¿Cuál es la probabilidad de que el tiempo entre dos fallas sucesivas en este equipo sea mayor que dos meses de trabajo?

Solución

Como $r(t)$ es un polinomio en t, la distribución del tiempo hasta la falla es la distribución de Weibull. Es necesario identificar los valores de α y β. Entonces,

$$2.7786t^{1.3} = \alpha\beta t^{\beta-1}.$$

Por consiguiente, $\beta - 1 = 1.3$, esto es $\beta = 2.3$. Como $\alpha\beta = 2.7786$, obtenemos

$$\alpha = 2.7786/\beta = 2.7786/2.3 = 1.208.$$

Se nos pide calcular $P\{T > 2\} = R(2)$. Sustituyendo $t = 2$ en la ecuación de $R(t)$, llegamos a

$$R(2) = e^{-1.208\times(2)^{2.3}} = 4.977 \times 10^{-4}.$$

Algunas veces ni las funciones de tasa de falla creciente ni decreciente describen bien las características de falla para un equipo determinado. Un caso característico es la función de tasa de falla en forma de "tina", como la que se ve en la figura 12-4. En las primeras fases de la vida del producto, la tasa de fallas decrece. Esto se debe a que los componentes defectuosos fallan rápidamente, haciendo que las tasas de falla iniciales sean altas. A esto se le suele llamar etapa de mortalidad infantil. Una vez eliminados los componentes malos, la frecuencia de

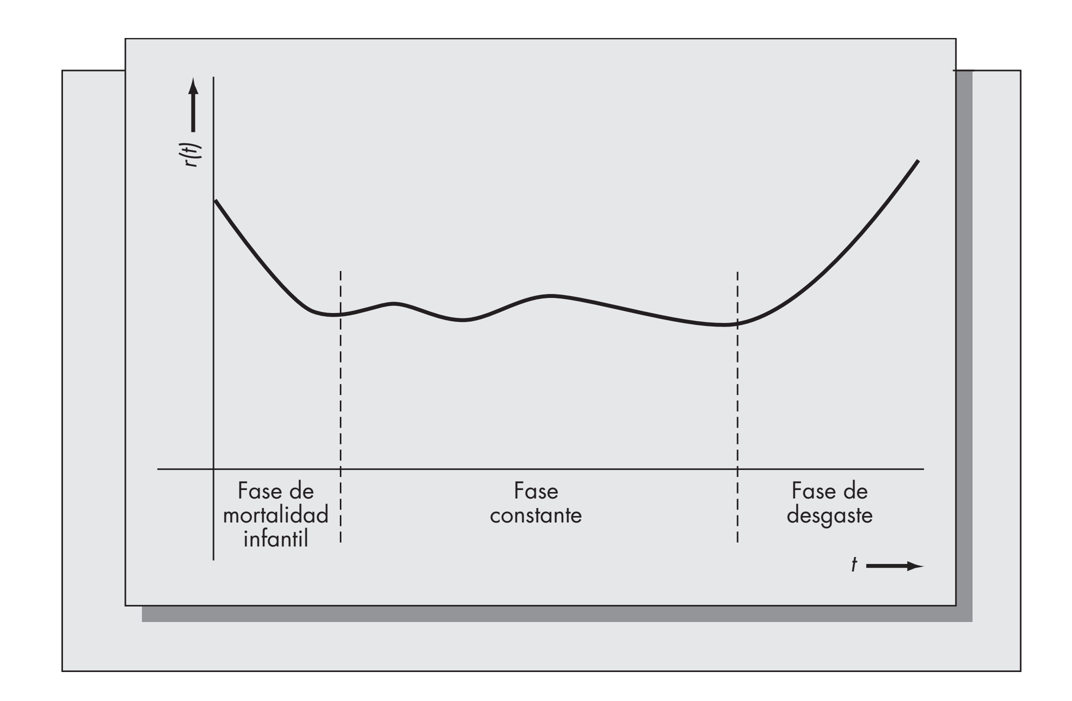


FIGURA 12-4
Función de tasa de falla en forma de "tina".

fallas permanece constante hasta que comienza el envejecimiento. En ese tiempo comienza la fase de desgaste, y la tasa de fallas comienza a aumentar.

Si $r(t)$ es una función arbitraria de tasa de fallas, se puede demostrar que la función de confiabilidad $R(t)$ se expresa como sigue:

$$R(t) = \exp\left(-\int_0^t r(u)\,du\right).$$

Es fácil comprobar que tanto el caso exponencial como el de Weibull satisfacen esta ecuación. A veces, como con la función de tasa de fallas en forma de tina, es difícil determinar una representación explícita de la función $r(t)$. En esos casos es posible aproximarla con funciones en escalera. Aquí no ilustraremos este procedimiento.

Problemas para la sección 12.2

7. Una pieza de equipo tiene una vida T (en años), variable aleatoria continua cuya función de distribución acumulada es

$$F(t) = 1 - e^{-t/10} - (t/10)\,e^{-t/10} \qquad \text{para toda } t \geq 0.$$

   a) ¿Cuál es la función de densidad de probabilidad de T?
   b) ¿Cuál es la probabilidad de que una parte sobreviva más de 20 años?
   c) ¿Cuál es la probabilidad de que una parte dure más de 10 años, pero menos de 20 años?
   d) ¿Cuál es la probabilidad de que una parte sobreviva más de 20 años, dado que ya ha durado 10 años?

8. Para el equipo del problema anterior,

   a) Deduzca la función $r(t)$, de tasa de fallas, y trace su gráfica correspondiente.
   b) Sin usar la función de tasa de fallas, determine la probabilidad de que una parte que ha durado 20 años de trabajo falle durante el año 21.
   c) ¿Determina $r(20)$ con exactitud la respuesta al inciso (*b*)? ¿Por qué sí o por qué no?

9. La Fuerza Aérea guarda cantidades enormes de datos sobre tiempos de falla de motores. Determinado motor ha tenido un patrón de fallas cuya función de tasa de fallas se aproxima bien mediante

$$r(t) = 0.000355e^{2.2t},$$

donde t está en horas de vuelo.

a) ¿Cuáles son las funciones de confiabilidad y de distribución acumulada del tiempo hasta la falla?

b) Determine el valor de t tal que la probabilidad de que falle un motor antes de t sea igual a la probabilidad de que falle después de t.

10. Una muestra de resistores de gran capacidad se prueba hasta la falla, y se ve que los resultados se ajustan a un modelo de probabilidades de Weibull. Con base en esos resultados, se estima que la función de confiabilidad de un resistor es

$$R(t) = e^{-0.0013\,t^{1.83}}.$$

a) ¿Cuál es la función de tasa de fallas para esos resistores?

b) ¿Es más probable que falle el resistor a medida que envejece?

c) ¿Cuál es la probabilidad de que un resistor funcione durante más de 30 horas sin que falle?

d) Suponga que un resistor ha estado funcionando durante 50 horas. ¿Cuál es la probabilidad de que falle en la hora 51? [Para sus cálculos, use los resultados del inciso *a*).]

12.3 EL PROCESO DE POISSON EN EL MODELADO DE LA CONFIABILIDAD

Considere una sola pieza de un equipo en funcionamiento que falla totalmente al azar. Como vimos en la sección anterior, eso indica que el tiempo hasta la falla tiene la distribución exponencial. Suponga que cuando falla la pieza, se repara de inmediato, o que el tiempo de reparación es suficientemente pequeño en comparación con el tiempo entre fallas y se puede ignorar. Así obtenemos un proceso en el que los eventos (fallas) se presentan al paso del tiempo, y en el que los tiempos entre fallas sucesivas son variables independientes exponenciales, idénticamente distribuidas. Sean T_1 $T_2, \ldots,$ variables aleatorias que corresponden a los tiempos entre fallas sucesivas. Cada una de las variables aleatorias tiene la función de distribución $F(t)$ y la función de confiabilidad $R(t)$ expresadas por

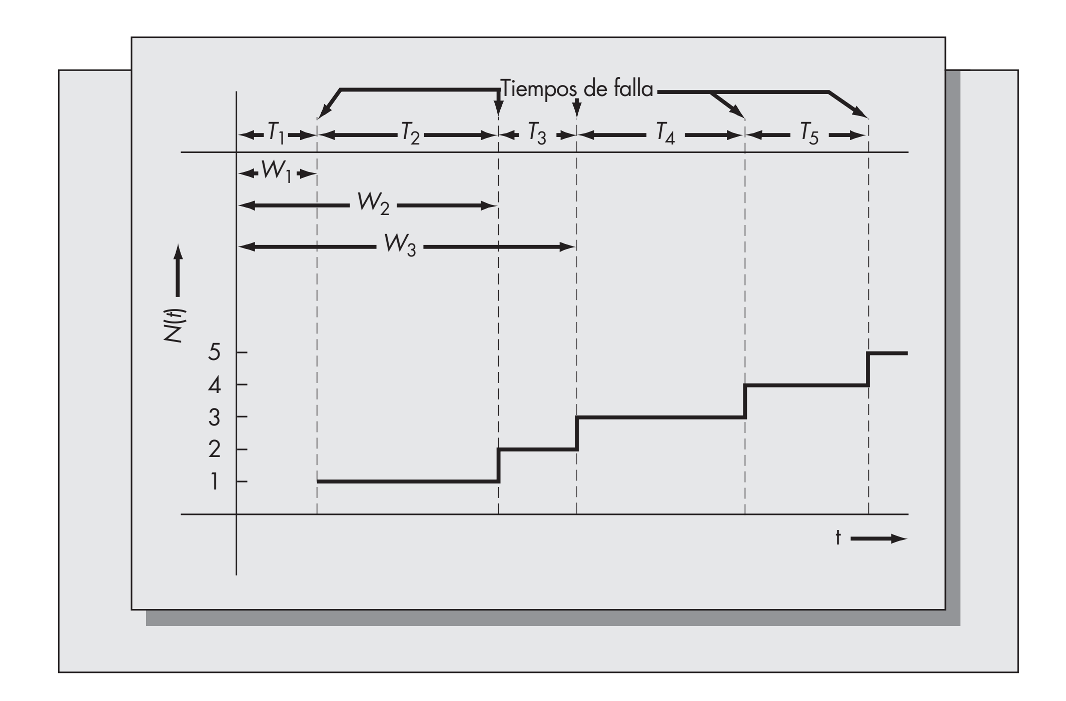
$$F(t) = 1 - e^{-\lambda t},$$
$$R(t) = e^{-\lambda t},$$

donde λ es la frecuencia con que ocurren las fallas. También definiremos una sucesión relacionada de variables aleatorias $W_1, W_2, \ldots,$ como sigue:

$$W_1 = T_1,$$
$$W_2 = T_1 + T_2,$$
$$W_3 = T_1 + T_2 + T_3,$$
$$\text{y así sucesivamente.}$$

Se interpreta que W_n es el tiempo de la n-ésima falla. Por último, introduciremos el proceso $N(t)$, que se define como la cantidad de fallas que se presentan hasta el tiempo t. $N(t)$ es un proceso *estocástico,* porque para cada valor fijo de t, $N(t)$ es una variable aleatoria. Es claro que

FIGURA 12-5
Realización de un proceso de Poisson.



$N(t)$ se relaciona estrechamente tanto con los tiempos entre fallas, $T_1, T_2, \ldots$, como con los tiempos de fallas $W_1, W_2, \ldots$. Cuando $T_1, T_2, \ldots$, son variables aleatorias independientes con distribución exponencial, $N(t)$ es un *proceso de Poisson.* En la figura 12-5 se muestra una realización de un proceso de Poisson.

Procederemos con el análisis del proceso de Poisson de la siguiente manera: se inicia conociendo la distribución de los tiempos entre fallas, $T_1, T_2, \ldots$. De ahí es posible deducir la distribución de los tiempos hasta la falla, $W_1, W_2, \ldots$. Se obtiene la distribución de $N(t)$ mediante la siguiente equivalencia de eventos:

$$\{N(t) < n\} = \{W_n > t\}.$$

Una pequeña reflexión puede convencernos de la validez de esta identidad. Para que sea cierto el lado izquierdo, la cantidad de fallas hasta el tiempo t debe ser menor que n. En ese caso, el tiempo de la n-ésima falla debe caer después de t, con lo que se obtiene el lado derecho. De igual forma, si la n-ésima falla se presenta después de t, la cantidad de fallas hasta el tiempo t debe ser menor que n.

El análisis requiere obtener la función de distribución acumulada de los tiempos hasta la falla. Como

$$W_n = T_1 + T_2 + \cdots + T_n,$$

se puede obtener la distribución de W_n formando la n-ésima convolución de la distribución exponencial, y esto conduce a la distribución *Erlang*:

$$P\{W_n > t\} = \sum_{k=0}^{n-1} \frac{e^{-\lambda t}(\lambda t)^k}{k!}.$$

La distribución lleva este nombre como reconocimiento al trabajo precursor de A. K. Erlang en el área de teoría de colas. (La distribución de Erlang para W_n se demuestra en el libro de Hillier y Lieberman, 1986, pp. 565-566.) Ahora podemos deducir $N(t)$, la distribución de la cantidad de fallas hasta el tiempo t.

TABLA 12-1 Resumen de resultados para el proceso de Poisson

Variable aleatoria	Distribución	Parámetro(s)	Media	Varianza
Tiempo entre fallas, T_n	Exponencial	λ	$1/\lambda$	$1/\lambda^2$
Tiempo de la n-ésima falla, W_n	Erlang	λ, n	n/λ	n/λ^2
Cantidad de fallas hasta el tiempo t, $N(t)$	Poisson	λt	λt	λt

Se tiene que

$$\begin{aligned} P\{N(t) = n\} &= P\{N(t) < n + 1\} - P\{N(t) < n\} \\ &= \sum_{k=0}^{n} \frac{e^{-\lambda t}(\lambda t)^k}{k!} - \sum_{k=0}^{n-1} \frac{e^{-\lambda t}(\lambda t)^k}{k!} \\ &= \frac{e^{-\lambda t}\,(\lambda t)^n}{n!}. \end{aligned}$$

Ésta es exactamente la distribución de Poisson con parámetro λt. La distribución de Poisson es discreta, y supone sólo valores de enteros no negativos 0, 1, 2, ... Tiene media y varianza dadas por λt. El proceso por el cual: (1) los eventos suceden totalmente al azar a través del tiempo, (2) los tiempos entre los eventos son variables aleatorias exponenciales, (3) los tiempos de los eventos son variables aleatorias con distribución de Erlang, y (4) la cantidad de eventos hasta cualquier tiempo t es una variable aleatoria con distribución Poisson, se llama proceso de Poisson. Es un proceso estocástico fundamental que surge en muchos campos, además de la teoría de la confiabilidad. El único parámetro λ, la tasa con la que suceden los eventos, es el que define el proceso. La tabla 12-1 es un resumen de los puntos importantes acerca del proceso de Poisson.

Como W_n es la suma de variables aleatorias independientes idénticamente distribuidas, cuando n es razonablemente grande, el teorema del límite central indica que se puede usar la distribución normal para aproximar la de Erlang. También, obsérvese que la función de confiabilidad de la distribución de Erlang tiene precisamente la misma forma que la de la distribución acumulada de Poisson. Por consiguiente, se puede usar una tabla de la distribución Poisson (tabla A-3 al final del libro) para obtener probabilidades de Erlang exactas. Además, obsérvese que para grandes valores de λt, la distribución normal es una aproximación adecuada a la de Poisson. (Sin embargo, no se debe usar la distribución normal para aproximar la exponencial.)

Ejemplo *12.5*

Una base militar local mantiene muchos equipos distintos. Uno de ellos es un aparato sensible de radar que detecta la incursión de aviones enemigos. Las descomposturas del instrumento se presentan totalmente al azar, con una tasa promedio de 3 fallas al año. Por lo general, el equipo se repara dentro del mismo día en el que falla. Calcule lo siguiente:

a) La probabilidad de que el tiempo entre dos fallas sucesivas sea menor que un mes.
b) La probabilidad de que haya exactamente cinco descomposturas en determinado año.
c) La probabilidad de que haya más de 15 fallas en un periodo de 4 años.
d) El tiempo medio para que sucedan 100 fallas.
e) La probabilidad de que la falla 25 suceda después de 10 años de funcionamiento.

Solución

Como en todos los problemas de probabilidad, se aconseja estimar de manera burda la respuesta antes de comenzar los cálculos formales. Esto servirá como comprobación de los cálculos.

a) Como hay tres fallas por año en promedio, habrá un promedio de 4 meses entre fallas. Por consiguiente, la probabilidad de que el tiempo entre dos fallas sucesivas sea menor que 1 mes, debe ser menor que .5.

Sea T el tiempo entre dos fallas sucesivas cualesquiera. Sabemos que T tiene la distribución exponencial con parámetro $\lambda = 3$ por año. Debemos estar seguros de expresar todas las unidades de tiempo en años para resolver este problema. Así, calculamos

$$P\{T < 1/12\} = 1 - \exp(-\lambda t) = 1 - \exp(-3/12) = .22.$$

b) En este caso contamos la cantidad de las descomposturas en determinado año, y así la distribución adecuada es la de Poisson, con parámetro $\lambda t = 3 \times 1 = 3$.

$$P\{N(1) = 5\} = \frac{e^{-3}3^5}{5!} = .1008.$$

c) Deseamos calcular $P\{N(4) > 15\}$, donde $N(4)$ tiene la distribución de Poisson con parámetro $\lambda t = 3 \times 4 = .12$. De la tabla A-3 obtenemos

$$P\{N(4) > 15\} = P\{N(4) \geq 16\} = .1556.$$

Esta probabilidad también se puede aproximar con la distribución normal. Para usar una aproximación normal, se determina la variable estandarizada Z restando la media y dividiendo entre la desviación estándar. Por consiguiente,

$$P\{N(4) > 15\} \approx P\{Z > (15 - 12)/\sqrt{12}\} = P\{Z > 0.8660\} = .1922.$$

Se puede mejorar esta aproximación aplicando la corrección por continuidad. Esa corrección es adecuada cuando se aproxima una variable aleatoria discreta mediante una variable continua. Como $\{N(4) > 15\} = \{N(4) \geq 16\}$, tomando un valor intermedio, usamos 15.5. (La corrección por continuidad se describe con detalle en el apéndice 11-A del capítulo 11.) Por consiguiente, $P\{N(4) > 15\} \approx P\{Z > (15.5 - 12)/\sqrt{12}\} = P\{Z > 1.01\} = .1562$. Es muy cercano a .1556, la respuesta exacta.

d) En este caso nos interesa $E(W_{100})$, de la tabla 12-1, $E(W_{100}) = 100/\lambda = 100/3 = 33.33$ años.

e) El tiempo de la falla 25 es la variable aleatoria W_{25}. De nuevo usaremos la aproximación normal, pero no necesitamos la corrección por continuidad, porque tanto la distribución de Erlang como la normal son continuas. De la tabla tenemos que $E(W_{25}) = 25/\lambda = 25/3 = 8.33$, y $\text{Var}(W_{25}) = 25/\lambda^2 = 25/9$. Por consiguiente, $\sigma = 5/3 = 1.67$. Entonces

$$P\{W_{25} > 10\} \approx P\left\{Z > \frac{10 - 8.33}{1.67}\right\} = P\{Z > 1\} = .1587.$$

(No se requiere la corrección por continuidad porque W_{25} es continua.)

Muchas aplicaciones implican la vigilancia de muchas piezas, y no sólo una de ellas. Por ejemplo, un reparador es el responsable de supervisar todo el equipo, y las aerolíneas deben mantener toda una flotilla de aviones. A continuación se describen las fallas de conjuntos de equipos.

Sistemas en serie sujetos a fallas puramente aleatorias

Considere un banco de artículos identificados por $1, 2, \ldots, N$, y supóngase que cada uno falla totalmente al azar, esto es, siguiendo una ley exponencial de fallas. Además, se supone que los artículos fallan en forma independiente. Un sistema en serie implica que el banco falla cuando falla el primer artículo en dicho sistema. Sean $T_1, T_2, \ldots, T_N$ los tiempos de falla asociados a cada pieza del equipo. Entonces

$$P\{T_i > t\} = \exp(-\lambda_i t) \qquad \text{para } 1 \leq i \leq N.$$

Definiremos a la variable aleatoria $T = \text{mín}(T_1, T_2, \ldots, T_N)$. Entonces T representa el tiempo en el que fallará el siguiente componente. También es el tiempo en el que falla el banco.

$$\begin{aligned} P\{T > t\} &= P\{\text{mín}(T_1, T_2, \ldots, T_N) > t\} \\ &= P\{T_1 > t, T_2 > t, \ldots, T_N > t\} \end{aligned}$$

(esto se deduce de que si el mínimo de un grupo de números es mayor que un número fijo, entonces todos los miembros del grupo también serán mayores).

$$= P\{T_1 > t\} \times P\{T_2 > t\} \times \cdots \times P\{T_N > t\}$$

(esto se debe a la independencia entre los tiempos individuales de falla)

$$= e^{-\lambda_1 t} e^{-\lambda_2 t} \cdots e^{-\lambda_N t}$$

$$= \exp\left(-\sum_{i=1}^{N} \lambda_i t\right),$$

que es exactamente la ley exponencial de fallas con $\lambda = \Sigma\, \lambda_i$. Si suponemos que las unidades que fallan se reparan con rapidez, la cantidad de fallas del banco hasta cualquier tiempo t, digamos $N(t)$, será un proceso de Poisson con tasa λ.

Problemas para la sección 12.3

11. Los automóviles llegan a una caseta de cobro de una autopista en forma totalmente aleatoria, siguiendo un proceso de Poisson con tasa $\lambda = 4$ automóviles por hora. Calcule lo siguiente:
    - *a*) La probabilidad de que el tiempo entre dos llegadas sucesivas cualesquiera sea mayor que 20 minutos.
    - *b*) La probabilidad de que lleguen exactamente cuatro automóviles en cualquier intervalo de una hora.
    - *c*) La probabilidad de que lleguen más de cinco automóviles en cualquier hora.
    - *d*) La probabilidad de que lleguen más de 10 automóviles en cualquier periodo de 2 horas.
    - *e*) Un empleado comienza a contar las llegadas a la caseta a las 8 A. M. ¿Cuál es la probabilidad de que cuente 20 llegadas antes de las 12, en el mediodía? (Para sus cálculos, use una aproximación normal.)
12. Herman's Hardware tiene un letrero de neón que está encendido siempre. Ese letrero tiene una vida promedio de 1 250 horas, y falla completamente al azar. Los letreros que se queman se remplazan al instante.
    - *a*) En promedio, ¿cuántos letreros de neón usa Herman's en un año?
    - *b*) Suponga que los letreros cuestan 37.50 dólares cada uno y que el señor Herman, el propietario, ha presupuestado 300 dólares anuales para remplazarlos. ¿Cuál es la probabilidad de que este presupuesto no alcance en cualquier año dado?
    - *c*) ¿Cuál es la probabilidad de usar dos letreros en el mismo mes? (Suponga, en sus cálculos, que un mes equivale a 30 días.)
13. Un módulo electrónico que usa la Marina en un instrumento de sonar requiere remplazarse, en promedio, una vez cada 16 meses, y falla de acuerdo con un proceso de Poisson. Suponga que la Marina pone en servicio equipos de sonar en la misma fecha, en ocho barcos distintos. Si los módulos se remplazan inmediatamente después de que fallan, y el presupuesto prevé exactamente 40 repuestos para 5 años, ¿cuál es la probabilidad de que el presupuesto sea insuficiente? (Sugerencia: use una aproximación normal a la de Poisson.)
14. Para el problema anterior, calcule lo siguiente:
    - *a*) La probabilidad de que un solo barco, en misión de 6 meses, no requiera remplazo del módulo durante su misión.
    - *b*) La probabilidad de que el tiempo de la quinta falla sea más de un año después de haberse puesto en funcionamiento los aparatos.
    - *c*) El tiempo esperado en el que se usarán los 40 repuestos.

12.4 FALLAS DE EQUIPO COMPLEJO

Muchas aplicaciones de la teoría de confiabilidad implican la predicción de los patrones de falla del equipo, según el conocimiento de los patrones de falla de los componentes de ese equipo. Por ejemplo, un conocido estudio que publicó la Comisión Reguladora Nuclear en Estados Unidos (1975) afirmaba que las plantas nucleares son seguras. En ese estudio se aplicó la teoría de confiabilidad para predecir qué tan probable sería un gran problema en una planta nuclear, analizando la tasa de falla de sus diversos componentes. Desafortunadamente, los supuestos erróneos sobre la independencia entre esos componentes condujeron a la conclusión de que los accidentes nucleares graves son virtualmente imposibles. Es obvio que no es así.

En la sección 12.3 demostramos que un banco de artículos conectados en serie, cada uno de los cuales está sujeto a la ley exponencial de fallas, también tendrá la ley exponencial de fallas. Ese resultado es un caso especial que deduciremos en esta sección.

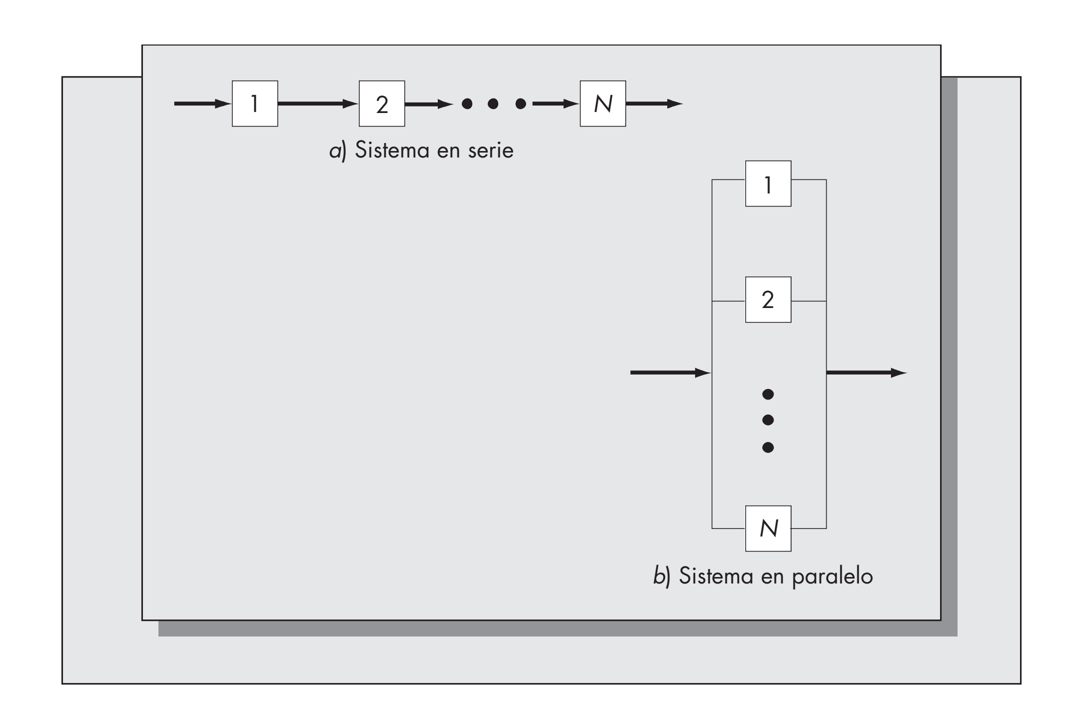
Componentes en serie

Un sistema en serie funciona sólo si cada componente lo hace. La figura 12-6*a* ilustra un diagrama esquemático de un sistema en serie.

Definiremos a T_i como el tiempo hasta la falla para el i-ésimo componente, siendo T_S el tiempo de falla de todo el sistema en serie. Como antes, $T_S = \text{mín}(T_1, T_2, \ldots, T_N)$. Recordemos la definición de la función de confiabilidad, $R(t) = P\{T > t\}$. Deduciremos $R_S(t)$, la función de confiabilidad del sistema, en términos de las funciones de confiabilidad de cada uno de los componentes, $R_i(t)$. Usando esencialmente los mismos argumentos que en la sección anterior,

$$\begin{aligned} R_S(t) &= P\{T_S > t\} = P\{\text{mín}(T_1, T_2, \ldots, T_N) > t\} \\ &= P\{T_1 > t, T_2 > t, \ldots, T_N > t\} \\ &= P\{T_1 > t\} \times P\{T_2 > t\} \times \cdots \times P\{T_N > t\} \\ &= R_1(t) \times R_2(t) \times \cdots \times R_N(t). \end{aligned}$$

FIGURA 12-6
Sistemas de componentes tanto en serie como en paralelo.



Para N componentes idénticos, cada uno con la función de confiabilidad $R(t)$, esto se convierte simplemente en $[R(t)]^N$.

Para expresar la función de distribución acumulada del sistema en serie, en términos de la función de distribución de cada uno de los componentes individuales, se sustituye $F(t) = 1 - R(t)$. Para N componentes idénticos en serie obtenemos

$$F_S(t) = 1 - [1 - F(t)]^N.$$

Componentes en paralelo

La figura 12-6*b* representa un sistema de componentes en paralelo. Un sistema en paralelo funciona si cualquiera de los componentes funciona. Estos sistemas se usan cuando se recurre a la redundancia para aumentar la confiabilidad. Definiremos a T_P como el tiempo de falla de un sistema en paralelo formado por componentes idénticos. Es claro que $T_P = \text{máx}(T_1, T_2, \ldots, T_N)$. Para un sistema en paralelo conviene más determinar la función de distribución del tiempo hasta la falla, y no la de la función de confiabilidad. Tenemos que

$$\begin{aligned} F_P(t) &= P\{\text{máx}(T_1, T_2, \ldots, T_N) \leq t\} \\ &= P\{T_1 \leq t, T_2 \leq t, \ldots, T_N \leq t\} \end{aligned}$$

(porque si el elemento más grande de un grupo es menor que determinado número, entonces todos los elementos del grupo también son menores que ese número)

$$= F_1(t) \times F_2(t) \times \cdots \times F_N(t),$$

que se reduce a $[F(t)]^N$ en el caso de N componentes idénticos.

La función de confiabilidad de un sistema con N componentes idénticos en paralelo es $R_P(t) = 1 - [1 - R(t)]^N$.

Cálculos de valor esperado

Un resultado útil de la teoría de la probabilidad, que puede simplificar el cálculo de los valores esperados, es el siguiente:

Teorema **Si T es una variable aleatoria no negativa con función de distribución acumulada $F(t)$ y función de densidad de probabilidad $f(t)$, entonces se puede calcular su valor esperado de dos maneras:**

$$E(T) = \int_0^\infty t f(t) dt = \int_0^\infty [1 - F(t)] dt.$$

La segunda ecuación puede agilizar los cálculos del tiempo esperado hasta la falla. Usaremos dicha ecuación para calcular el de un sistema de N componentes idénticos, cada uno de los cuales tiene la ley exponencial de fallas con parámetro λ.

Tenemos que

$$E(T_P) = \int_0^\infty [1 - (1 - e^{-\lambda t})^N] dt.$$

Para hacer la integración realizamos el cambio de variable $v = 1 - e^{-\lambda t}$, y se obtiene $dv = \lambda e^{-\lambda t} dt$, es decir

$$dt = \frac{1}{\lambda} \frac{1}{e^{-\lambda t}} dv = \frac{1}{\lambda} \frac{1}{1 - v} dv.$$

Por lo tanto,

$$E(T_P) = \frac{1}{\lambda} \int_0^1 \frac{1 - v^N}{1 - v} dv.$$

La expresión en el integrando es justamente la serie geométrica finita $1 + v + v^2 + \cdots + v^{N-1}$. Por consiguiente, se concluye que

$$E(T_P) = \frac{1}{\lambda}\int_0^1 (1 + v + v^2 + \cdots + v^{N-1})\,dv$$
$$= \frac{1}{\lambda}(1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + 1/n).$$

Esto implica que, para sistemas con k componentes en paralelo, la vida esperada del sistema aumenta en $1/(k + 1)\lambda$ cuando se agrega un componente más.

Ejemplo *12.6*

Wizard, una conocida marca de controlador eléctrico para puertas de cochera, incluye con cada uno de sus equipos dos focos de 40 watts que encienden cuando se abre la puerta. Por lo general, un bulbo dura más o menos un año en operación normal. Tres vecinos, James, Smith y Walker, tienen controladores Wizard en sus cocheras. Cada vez que se funde un foco, James remplaza los dos. Por otro lado, Smith sólo cambia el foco que se quemó, y Walker cambia los dos sólo cuando se ha quemado el segundo. Suponga que los focos fallan apegándose a la ley exponencial.

a) En un periodo de 10 años, ¿cuántos focos, en promedio, necesitará cada vecino?

b) ¿En qué porcentaje del tiempo Walker tendrá sólo un foco quemado?

c) ¿Hay alguna ventaja en la estrategia de James respecto a la de Smith?

Solución

a) Tanto James como Smith consideran éste como un sistema puro en serie de dos componentes: el sistema falla cuando uno de los focos falla. Si cada foco tiene una tasa de fallas $\lambda = 1$, entonces el sistema tiene una tasa de fallas $\lambda = \lambda_1 + \lambda_2 = 2$. Como James remplaza dos focos cada vez que hay una falla, usa un promedio de cuatro focos cada año, o 40 en 10 años. Por otra parte, Smith sólo requiere un foco cada vez que el sistema falla, así que usa un promedio de dos focos por año, o 20 cada 10 años.

La política de Walker, consistente en remplazar ambos focos sólo después de que ambos fallan, equivale a un sistema paralelo puro de dos componentes. La vida esperada de un sistema de dos componentes en paralelo con tasa de fallas 1 es $1 + \frac{1}{2} = 1.5$ años. Por consiguiente, cada 1.5 años necesita dos focos, y así se obtiene un promedio de 6.67 remplazos durante los 10 años, que equivale a 13.33 focos.

b) Cabría pensar que la mitad del tiempo tendría un foco funcionando y la otra mitad tendría dos. Sin embargo, no es así. Como la tasa de falla del sistema en serie es $\lambda = 2$, el primer foco fallará, en promedio, a los 6 meses. Como el sistema en paralelo tiene, en promedio, una vida de 1.5 años, el foco restante durará un promedio de un año. Por consiguiente, estará trabajando su sistema con un solo foco durante el 66.67% del tiempo en promedio.

c) Como la ley de falla es exponencial, la estrategia de James no aventaja en absoluto la de Smith. Ambos tendrán que hacer remplazos con la misma frecuencia (dos veces al año), pero James usará el doble de focos. Sin embargo, si la ley de falla no fuera exponencial, es posible que con el método de James se tuvieran menos remplazos de focos.

Este problema hace surgir un punto interesante. En instalaciones en las que se usan muchos focos, como en los hoteles de Las Vegas, se acostumbra una estrategia consistente en remplazar periódicamente todos los focos. Sin embargo, como vimos en el problema, si tienen una ley exponencial de fallas, esta estrategia no dará como resultado menores remplazos no planeados, y sí conducirá a usar mucho más focos.

Sistemas *K* de *N*

Supongamos que un sistema está formado por N componentes. Un sistema K de N es aquel que sólo funciona si trabajan cuando menos K de los componentes, donde $1 \le K \le N$. Un ejemplo característico de sistema K de N es el de un avión cuatrimotor, que puede volar cuando trabajan al menos dos de sus motores.

Para analizar un sistema K de N se usa un marco binomial. Imaginemos que cada componente es un ensayo separado de Bernoulli: un éxito es un componente que funciona, y un fracaso es uno que no funciona. Supongamos que todos los componentes son idénticos, y entonces cada uno tiene la misma función de confiabilidad, $R(t)$, y la misma función de distribución, $F(t)$. Fijemos un punto t en el tiempo. Sea $p = P\{\text{un componente funciona en el momento } t\} = R(t)$. La probabilidad $R_k(t)$ de que el sistema funcione en el tiempo t es la probabilidad de que haya al menos K éxitos en N intentos de un experimento binomial, con $p = P\{\text{éxito}\}$ en cada intento.

Por lo tanto,

$$R_K(t) = \sum_{j=K}^{N} \binom{N}{j} p^j (1-p)^{N-j}$$

$$= \sum_{j=K}^{N} \binom{N}{j} R(t)^j F(t)^{N-j}.$$

Observamos que ambos sistemas en serie y en paralelo son casos especiales de los sistemas K de N. Un sistema en serie es un sistema N de N, y un sistema en paralelo es un sistema 1 de N. Un sistema en serie de N componentes idénticos siempre tendrá menor confiabilidad que un sólo componente, mientras que un sistema de N componentes idénticos en paralelo siempre tendrá mayor confiabilidad que un sólo componente.

Sin embargo, el que un sistema K de N sea más confiable que un sólo componente depende de la confiabilidad de los componentes individuales. La figura 12-7 muestra una gráfica de la confiabilidad de sistemas 2 de 4 y 3 de 4, en función de la confiabilidad de cada componente. La línea de 45° representa la confiabilidad de un solo componente. La curva de confiabilidad para el sistema 2 de 4 cruza la recta de 45º aproximadamente en $p = .23$, y la del sistema 3 de 4 aproximadamente en $p = .77$. Esto indica que se prefiere un sistema 2 de 4 sobre uno de un solo componente, si la confiabilidad de los componentes que forman el sistema es mayor que .23, y que se prefiere un sistema 3 de 4 a uno de un solo componente sólo si la confiabilidad es mayor que .77.

FIGURA 12-7
Confiabilidad de los sistemas 2 de 4 y 3 de 4.

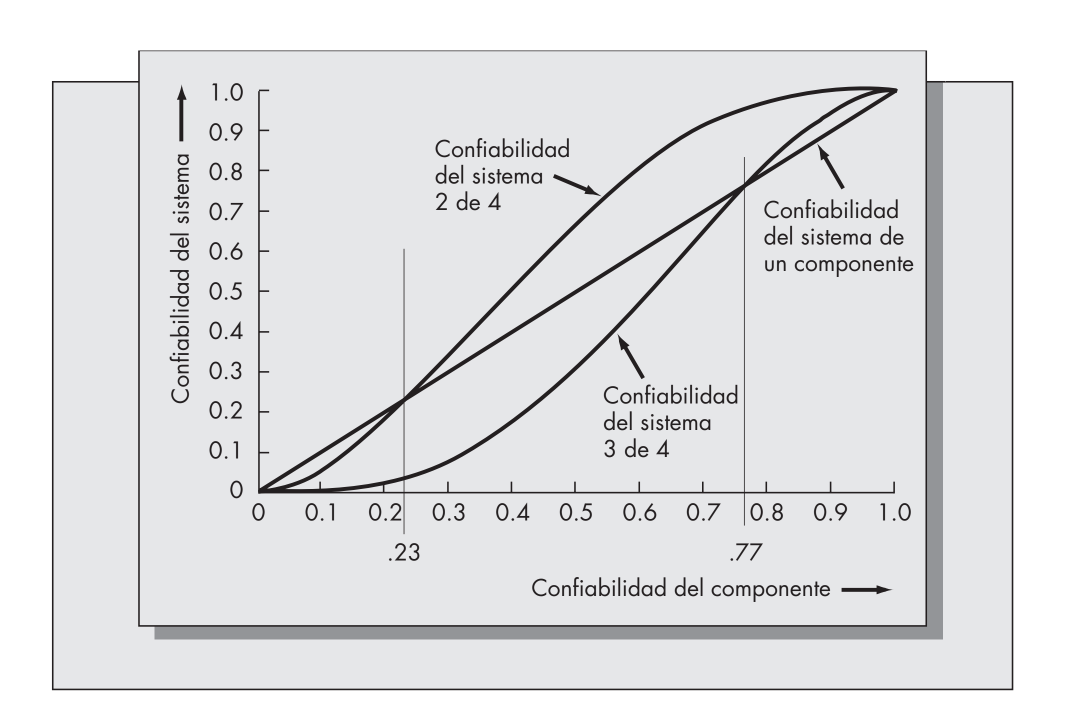
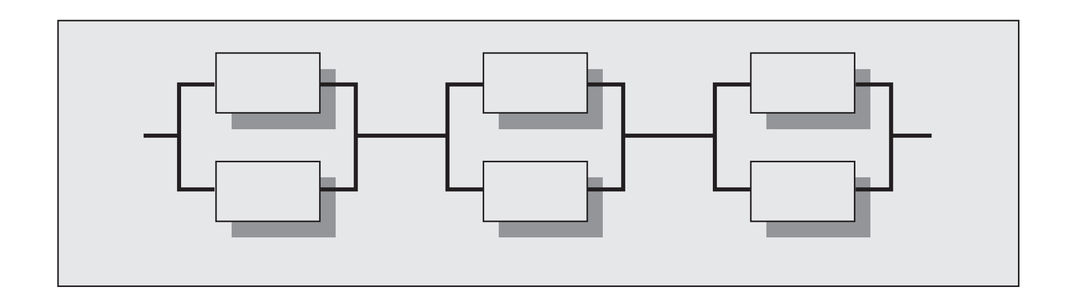


FIGURA 12-8
Sistema de componentes (para el problema 19).



Problemas para la sección 12.4

15. Un subensamble de un robot industrial está formado por 12 componentes en serie, y cada uno de ellos falla completamente al azar, con una frecuencia de una vez cada 50 años.

    a) ¿Cuál es el tiempo promedio entre fallas del robot?

    b) ¿Cuál es la probabilidad de que este subensamble no falle a los 8 años de funcionamiento?

16. Demuestre que el tiempo en que falla un sistema de n componentes idénticos en serie, cada uno de los cuales tiene una distribución Weibull para el tiempo de vida, también tiene distribución del tipo Weibull.
17. Considere los tres sistemas siguientes: *a*) componente único que tiene tasa de fallas igual a 1 por año, *b*) dos componentes en serie, cada uno de los cuales tiene tasa de fallas igual a 1 cada 2 años, y *c*) dos componentes en paralelo, cada uno de los cuales tiene tasa de fallas igual a 2 por año. Compare la confiabilidad de esos tres sistemas, suponiendo una ley exponencial de fallas.
18. Un diseñador quiere determinar la cantidad de niveles de redundancia que incluirá en determinado circuito. Este circuito será parte de una pieza sensible cuyo costo se estima en 500 dólares por falla. Cada nivel adicional de redundancia cuesta 100 dólares. Si cada componente falla al azar, con una tasa igual a una falla cada 5 años, ¿qué nivel de redundancia se acerca más estrechamente al costo del diseño, con el costo esperado de falla del equipo a lo largo de su ciclo de vida de 10 años?
19. Considere el sistema de seis componentes idénticos de la figura 12-8. Si cada componente tiene una tasa constante de falla igual a λ, deduzca la función de distribución del tiempo hasta la falla del sistema.
20. Un motor de avión falla con la probabilidad p. Suponga que para que un avión termine bien un vuelo, deben trabajar cuando menos la mitad de los motores. Demuestre que para $0 < p < \frac{1}{3}$, es preferible un cuatrimotor a un bimotor, y que para $\frac{1}{3} < p < 1$, un bimotor es preferible a un cuatrimotor.

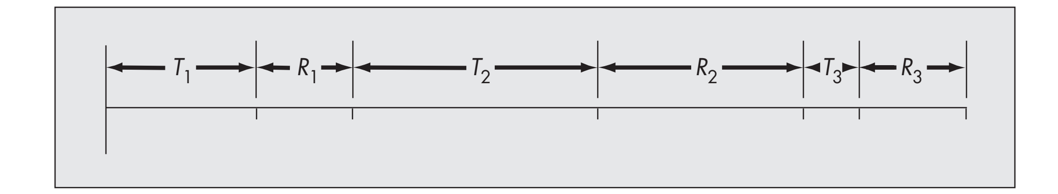
12.5 INTRODUCCIÓN A LOS MODELOS DE MANTENIMIENTO

Con frecuencia el mantenimiento de equipos complejos puede consumir una gran parte de los costos asociados con esos equipos. Por ejemplo, se ha estimado que los costos de mantenimiento en el área militar son casi la tercera parte de todos los costos de operación en que se incurre. Es claro que los temas de confiabilidad y de mantenimiento se relacionan mucho.

En esta sección introduciremos cierta terminología normal en el mantenimiento:

1. TMEF = tiempo medio entre fallas. Corresponde al tiempo esperado entre fallas en nuestra notación anterior, y es igual a $1/\lambda$.
2. TPDR = tiempo medio de reparación. Es el valor esperado del tiempo de reparación R.

FIGURA 12-9
Realización de tiempos de falla y reparación.



3. Disponibilidad = Fracción promedio del tiempo que trabaja el equipo. Está dada por la fórmula

$$\text{Disponibilidad} = \frac{E(T_i)}{E(T_i) + E(R_i)} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}.$$

Podemos pensar que una pieza única de equipo tiene tiempos sucesivos de falla T_1, T_2, . . . , y tiempos sucesivos de reparación R_1, R_2, ... La figura 12-9 muestra un diagrama esquemático de ese sistema.

Ejemplo *12.7* El tiempo medio entre fallas de una máquina copiadora es 400 horas de funcionamiento. Las reparaciones requieren normalmente una media de 10 horas, desde que se recibe la solicitud de reparación hasta que se termina el servicio. Calcule la disponibilidad de esta copiadora.

Solución La disponibilidad es 400/(400 + 10) = 400/410 = .9756.

El ciclo de reparación se define como el tiempo entre dos reparaciones sucesivas. Con frecuencia debemos determinar la distribución de un ciclo de reparación, más que sólo su esperanza. Un solo ciclo de reparación es la suma del tiempo hasta la falla T_1 y el tiempo de reparación R_1. La distribución exacta de la suma de dos variables aleatorias es la *convolución* de las distribuciones individuales. (Por ejemplo, véase DeGroot, 1986.)

Si los tiempos entre fallas y de reparación se pueden aproximar en forma razonable con la distribución normal, la suma de los dos también tendrá distribución aproximadamente normal.

Ejemplo *12.8* Suponga que el tiempo entre fallas, T_i, tiene distribución aproximadamente normal con media de 400 horas y varianza de 10 000. El tiempo de reparación del equipo también tiene distribución aproximadamente normal, con media de 10 horas y varianza de 11.6. Calcule la probabilidad de que haya más de seis ciclos de reparación en un periodo de un año. Suponga que un año equivale a 2 000 horas de operación.

Solución Tenemos que,

$$E(T_i + R_i) = 400 + 10 = 410,$$
$$\text{Var}(T_i + R_i) = 10\,000 + 11.6 = 10\,011.6.$$

De aquí se deduce que,

$$E\left[\sum_{i=1}^{6} (T_i + R_i)\right] = 6 \times 410 = 2\,460,$$

$$\text{Var}\left[\sum_{i=1}^{6} (T_i + R_i)\right] = 6 \times 10\,011.6 = 60\,069.6,$$

$$P\left\{\sum_{i=1}^{6} (T_i + R_i) \leq 2\,000\right\} = P\left\{Z \leq \frac{2\,000 - 2\,460}{\sqrt{60\,069.6}}\right\}$$
$$= P\{Z \leq -1.88\} = .03.$$

12.6 ESTRATEGIAS DETERMINISTAS DE REMPLAZO POR EDAD

Para un equipo en operación que no sigue una ley exponencial de fallas, con frecuencia hay ventajas al remplazarlo *antes* de que falle. Esto es cierto cuando el costo de la reparación es mucho mayor si falla mientras está trabajando. En algunos casos, como en las operaciones militares, podría ser que una falla de equipo fuera imposible de corregir, y ocasionar que se pierdan vidas.

En esta sección consideraremos modelos de remplazo por edad que no consideran en forma explícita la incertidumbre del proceso de falla. Más bien, el mecanismo de envejecimiento se incluye en la estructura de costos en particular; se supone que el costo de mantener el equipo aumenta a medida que envejece. En la sección siguiente se examinarán modelos de remplazo planeado, que incluyen en forma explícita la incertidumbre del proceso de falla.

Los modelos que veremos son adecuados para equipo con funcionamiento continuo, como los de radar o las centrales generadoras, y para equipo con funcionamiento intermitente, como los automóviles. En el último caso se registraría el tiempo de funcionamiento, no el tiempo total transcurrido.

Supongamos que el costo de remplazo del artículo es K. También supondremos que la tasa instantánea del costo de trabajar un artículo de edad u es $C(u)$. Consideraremos varias formas de $C(u)$, pero al principio supondremos que $C(u) = au$.

Con base en los valores de los diversos costos, habrá un punto óptimo en el que debe remplazarse un artículo para minimizar el costo total por unidad de tiempo. La función del costo total puede considerarse como la suma de dos componentes: mantenimiento y remplazo. Conforme pasa el tiempo, el costo marginal del mantenimiento aumenta, y el costo marginal de remplazo disminuye. La edad óptima de remplazo minimiza la función de costo promedio. En los casos que se describirán, la función de costo promedio es convexa, por lo que será fácil determinar la solución óptima.

La política óptima en el caso básico

Haremos los siguientes supuestos:

1. El equipo que se usa trabaja continuamente.
2. No se toma en cuenta el tiempo muerto para reparación y mantenimiento.
3. El horizonte de planeación es infinito.
4. Cada equipo o pieza nueva tiene características idénticas.
5. Sólo se consideran costos de mantenimiento y de remplazo.
6. El objetivo es minimizar los costos de remplazo y mantenimiento a largo plazo.
7. La tasa de costo de mantener un artículo de edad u es au, y el costo de remplazar el artículo es K. No hay valor de recuperación o salvamento.

La variable de decisión es el tiempo que transcurre desde el punto en que se compra un equipo o parte hasta que se remplaza por un artículo nuevo. La figura 12-10 muestra ciclos sucesivos de remplazo.

El objeto del análisis es determinar el valor de t que minimice el costo total de mantenimiento y remplazo durante un horizonte infinito. Un ciclo de remplazo es el tiempo entre remplazos sucesivos. Como todos los ciclos de remplazo son idénticos, sólo pondremos atención a los costos incurridos en un solo ciclo.

$$\text{Costo total de remplazo por ciclo} = K,$$

$$\text{Costos totales de mantenimiento por ciclo} = \int_0^t C(u)\,du = \int_0^t au\,du = \frac{at^2}{2}.$$

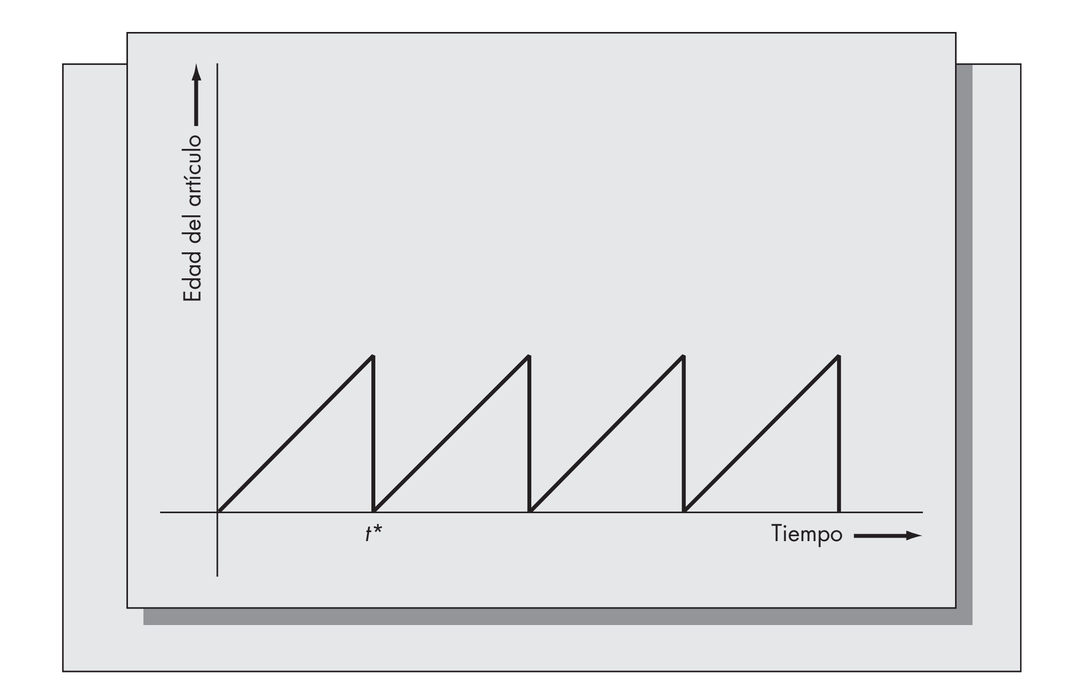


FIGURA 12-10
Estrategia óptima de remplazo por edad.

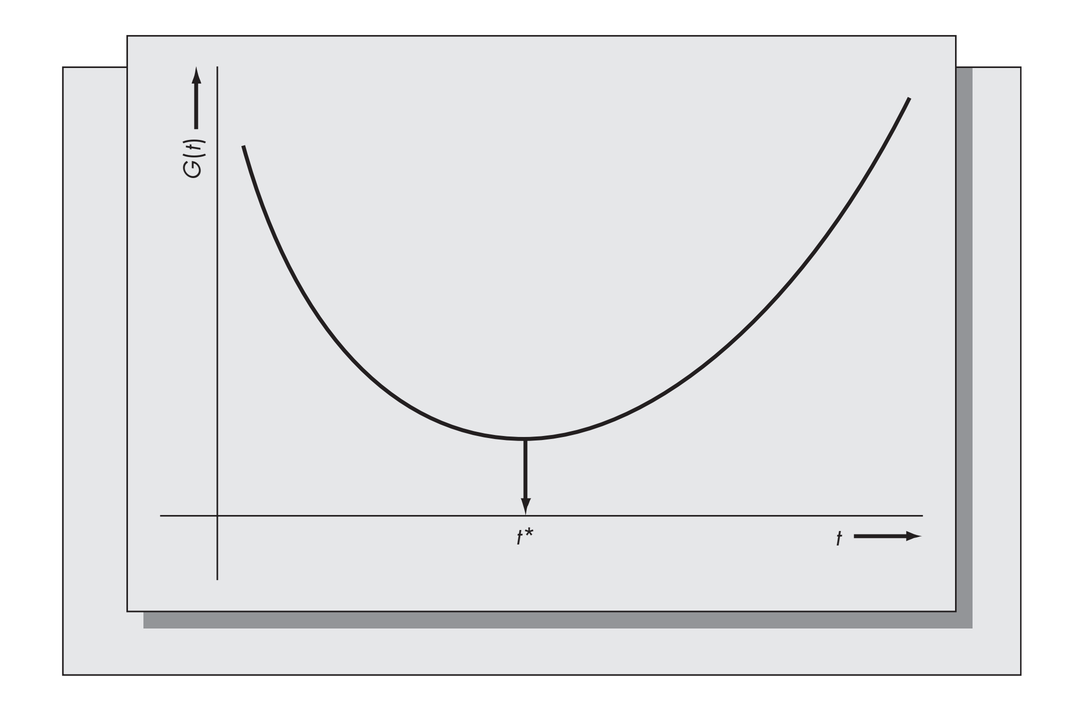


FIGURA 12-11
Cálculo de la edad de remplazo óptima.

El costo promedio por unidad de tiempo es justamente el costo total del ciclo dividido entre la longitud del ciclo. Sea $G(t)$ el costo promedio por unidad de tiempo, cuando el tiempo de remplazo es t. En ese caso,

$$G(t) = \frac{1}{t}\left(K + \frac{at^2}{2}\right) = \frac{K}{t} + \frac{at}{2}.$$

Como $G''(t) = K/t^3 > 0$, entonces G(t) es una función convexa de la variable única t. La función $G(t)$ se muestra en la figura 12-11. El objetivo es determinar el valor óptimo de t, digamos t^*, que minimiza a $G(t)$. Como $G(t)$ es convexa, la solución óptima satisface a

$$G'(t) = \frac{-K}{t^2} + \frac{a}{2} = 0,$$

que resulta en

$$t^* = \sqrt{\frac{2k}{}}.$$

Ejemplo *12.9* Usaremos la versión sencilla del modelo de remplazo por edad para estimar la cantidad de años que debe conservarse un automóvil. Aunque el modelo no describe fielmente el problema del remplazo del automóvil, lo podemos usar como aproximación. Supongamos que el costo de mantenimiento anual de un vehículo con u años de antigüedad es $400u$ dólares. Esto quiere decir que el costo de mantenimiento durante el primer año es 200 dólares, durante el segundo es 600 dólares, durante el tercero es 1 000 dólares, y así sucesivamente. (Estas cantidades se obtienen calculando $at^2/2$ para $t = 1$, 2 y 3, y restando los costos de los años anteriores.) Es probable que los costos suban así con mayor rapidez que los costos reales de mantenimiento. Supongamos que un automóvil nuevo cuesta 10 000 dólares. De acuerdo con la fórmula, la cantidad óptima de años que debe conservarse el automóvil es

$$t^* = \sqrt{\frac{2K}{a}} = \sqrt{\frac{(2)(10\,000)}{400}} = 7.07 \text{ años}.$$

Un modelo general de remplazo por edad

Cuando incluimos un valor de salvamento, y permitimos una función costo de mantenimiento más general, el modelo se complica mucho. Supondremos *como* antes, que $C(u)$ es una función arbitraria que representa la tasa del costo de mantener un artículo cuya edad es u. Definimos $S(u)$ *como* el valor de salvamento de un artículo de edad u. Entonces el costo total incurrido en el ciclo es

$$K + \int_0^t C(u)\,du - S(t).$$

El costo promedio por unidad de tiempo es

$$G(t) = \frac{K}{} + \frac{1}{} \int_0^t C(u)\,du - \frac{S(t)}{}.$$

El valor óptimo de t, que es t^*, es la solución de

$$G'(t) = \frac{-K}{t^2} + \frac{H(t)}{t^2} + \frac{C(t)}{} + \frac{S(t)}{t^2} - \frac{S'(t)}{} = 0,$$

o bien

$$tC(t) + S(t) = K + H(t) + tS'(t),$$

donde por conveniencia definimos

$$H(t) = \int_0^t C(u)\,du.$$

La determinación de t^* puede ser bastante difícil. [Por ejemplo, trate de obtener una solución suponiendo que $C(u) = au$ y $S(u) = K - bu$.] Para muchos problemas reales, la distribución exponencial proporciona una descripción fiel del aumento en los costos de mantenimiento y la disminución del valor de reventa del equipo en funcionamiento. Si hacemos que

$$C(u) = ae^{bu}, \qquad \text{donde } a, b > 0,$$

y

$$S(u) = ce^{-du}, \qquad \text{donde } c, d > 0,$$

entonces el valor óptimo de t satisface a

$$tae^{bt} + ce^{-dt} = K + \int_0^t ae^{bu}\,du + t\frac{d}{}(ce^{-dt}).$$

Es fácil demostrar que

$$H(t) = \int_0^t ae^{bu}\,du = \frac{a}{}(e^{bt} - 1)$$

y que

$$\frac{d}{}(ce^{-dt}) = -cde^{-dt},$$

de modo que la ecuación que define a una solución óptima es

$$tae^{bt} + ce^{-dt} = K + \frac{a}{}(e^{bt} - 1) - tcde^{-dt}.$$

Reordenando los términos se obtiene

$$ae^{bt}\left(t - \frac{1}{}\right) + ce^{-dt}(1 + dt) + \frac{a}{} = K.$$

El objetivo es determinar el valor de t que acerca el lado izquierdo de la ecuación a K, el valor de remplazo, tan próximo como sea posible. Es una ecuación que se resuelve con dificultad para t porque implica exponenciales y constantes a la vez.[2] Con hojas de cálculo se puede obtener una solución. Tan sólo se calcula el lado izquierdo de la ecuación para diversos valores de t y se determina gráficamente el punto en el que esta función cruza el valor de K.

Ejemplo *12.10*

De nuevo intentaremos determinar el tiempo óptimo para cambiar un automóvil. Las funciones exponenciales son más realistas que las lineales, y deberían llegar a un estimado exacto del tiempo óptimo de remplazo. Como en el ejemplo 12.9, supondremos que el costo de remplazo del automóvil es 10 000 dólares, y que además se llama ecuación trascendente. El automóvil pierde 15% de su valor cada año, lo cual probablemente sea un estimado razonable de la declinación en el valor de reventa para la mayor parte de los automóviles nuevos. Esto significa que el automóvil vale (.85)(10 000) = $8 500 después de un año (.85)(.85)(10 000) = 7 225 después de 2 años, y así sucesivamente.

Se desea determinar c y d para que $S(t) = ce^{-dt}$ concuerde con estos valores. Como el valor de recuperación cuando el tiempo es $t = 0$ es exactamente el costo de remplazo, entonces $S(0) = 10\,000$. Al sustituir obtenemos.

$$S(0) = ce^{-d(0)} = c = 10\,000.$$

El valor del automóvil después de un año es (.85)(10 000), que corresponde a $S(1)$. Por lo tanto,

$$S(1) = (0.85)(10\,000) = ce^{-d(1)} = ce^{-d}.$$

Como $c = 10\,000$, obtenemos

$$e^{-d} = 0.85,$$
$$d = -\ln(0.85) = 0.1625.$$

[2] Se denomina una ecuación trascendental.

En consecuencia, se deduce que

$$S(t) = 10\,000e^{-0.1625t}.$$

Ahora consideremos los costos de mantenimiento. Supongamos, como en el ejemplo 12.9, que tales costos son de 200 dólares durante el primer año de funcionamiento. Esto equivale a

$$H(1) = 200$$

o bien

$$(a/b)(e^b - 1) = 200.$$

Además, supongamos que los costos de mantenimiento aumentan a la tasa de 40 por ciento por año. Esto quiere decir que

$$\frac{C(t)}{C(t-1)} = 1.4.$$

Sustituyendo $C(t)$ se obtiene

$$\frac{ae^{bt}}{ae^{b(t-1)}} = e^b = 1.40,$$

o $b = \ln(1.4) = .3365$. Se deduce que,

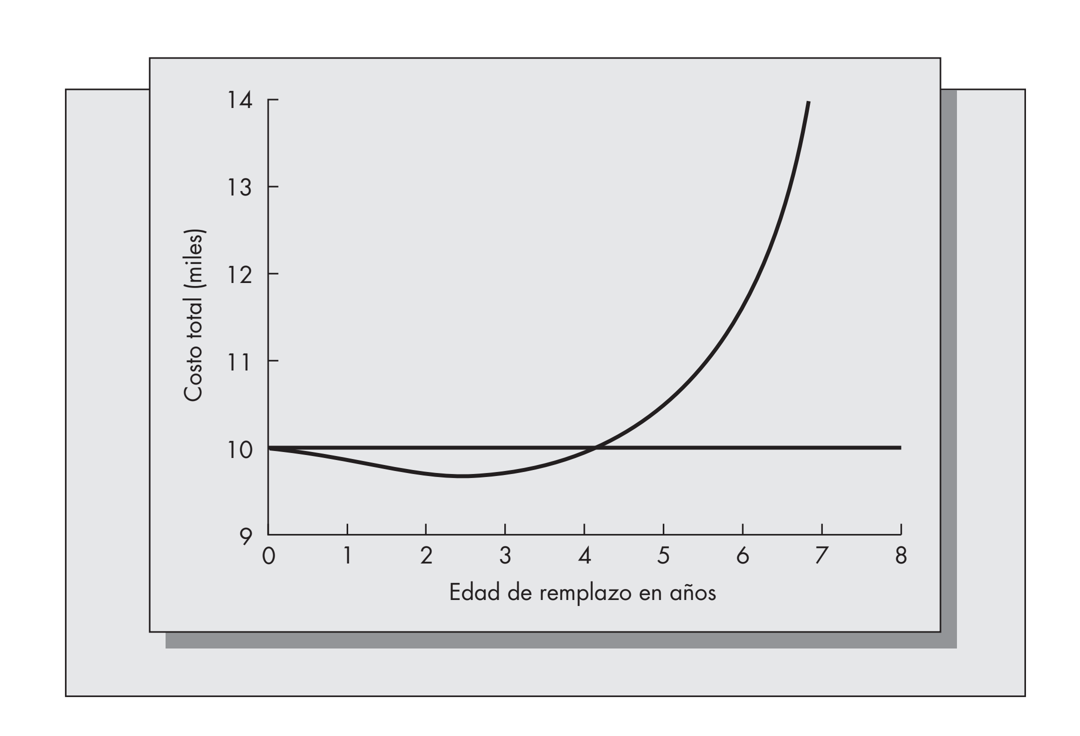
$$a = \frac{(200)(b)}{} = \frac{(200)(0.3365)}{0.4} = 168.25.$$

Combinando estos resultados, el tiempo óptimo para remplazar el automóvil, t^*, es el valor de t que resuelve la ecuación

$$168.25e^{0.3365t}(t - 2.972) + 10\,000e^{-0.1625t}(1 + 0.1625t) + 500 = 10\,000.$$

Un método para estimar la solución de esta ecuación es graficar la función representada por el lado izquierdo. Así se hizo en la figura 12-12. La edad de remplazo con costo mínimo es 4.5 años,

FIGURA 12-12
Número óptimo de años para remplazar un automóvil ($M = .4$).



aproximadamente. Esto significa que para un automóvil que cuesta 10 000 dólares cuyo valor disminuye al 15% anual y cuyo costo de mantenimiento es 200 dólares durante el primer año y aumenta el 40% anual, la estrategia óptima, que minimiza los costos promedio de remplazo, menos recuperación y más mantenimiento, es remplazarlo más o menos una vez cada cuatro años y medio.

Para ver el efecto del costo de mantenimiento hemos resuelto de nuevo el ejemplo con un aumento de 20%, en vez de 40% anual en costos de mantenimiento. Probablemente esto sea más exacto para la mayor parte de los vehículos. La solución se representa gráficamente en la figura 12-13. Aquí, el tiempo óptimo de remplazo es 9 años, aproximadamente.

Sea M la tasa con la que aumentan los costos de mantenimiento cada año, expresada como una fracción (en la figura 12-13, $M = .20$). Sea I_0 el costo de mantenimiento en el primer año, y D la tasa anual de depreciación, también expresada como fracción ($D = 0.15$ en el ejemplo). Entonces, puede demostrarse que

$$b = \ln(1 + M),$$

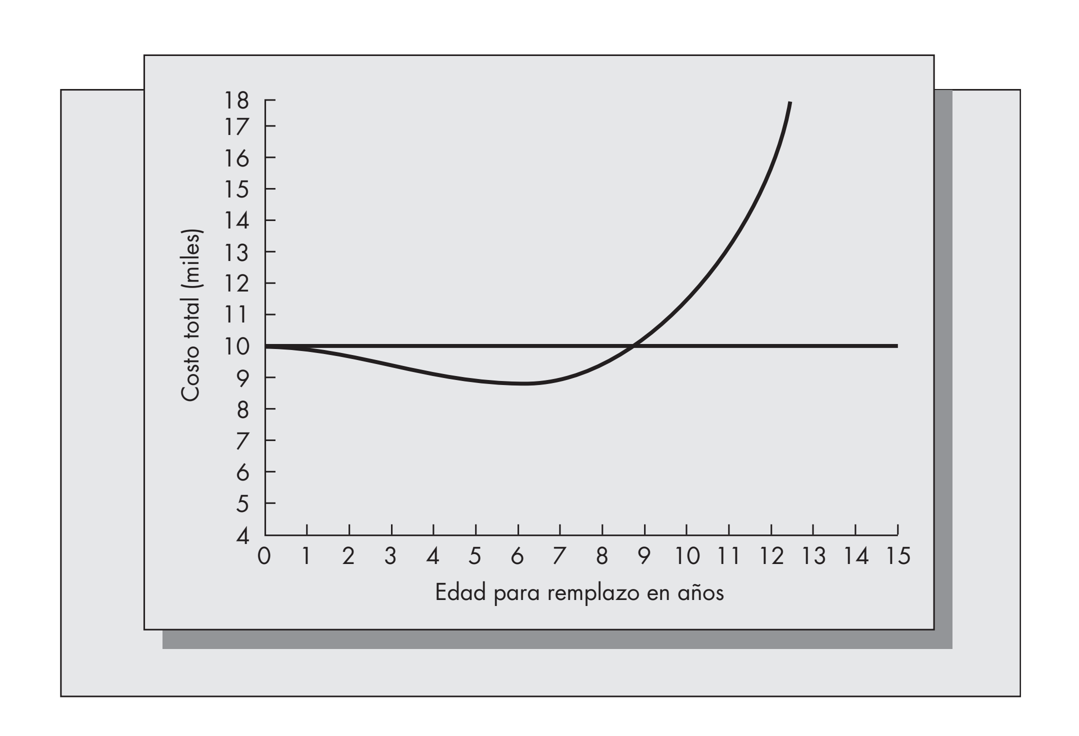
$$a = I_0 b/M,$$

$$c = K,$$

$$d = \ln(1 - D).$$

Los modelos presentados en esta sección no consideran los efectos de la inflación. Si se incluyeran, las soluciones obtenidas no cambiarían en forma apreciable, porque la inflación se aplica a los costos de remplazo y de mantenimiento por igual. El método es calcular el valor presente de todos los costos futuros descontados y determinar la estrategia de remplazo que minimice la función resultante. En esta sección no se describirán los flujos de caja descontados.

FIGURA 12-13
Número óptimo de años para remplazar un automóvil ($M = .2$).



Problemas para la sección 12.6

21. Para el modelo básico de remplazo por edad, imagine un equipo cuyo remplazo cuesta \$18 000. Los costos *totales* de mantenimiento durante 5 años de operación se estiman en \$2400. Suponiendo que los costos de mantenimiento aumentan linealmente, calcule el valor de a y la edad óptima a la que se debe remplazar el equipo.
22. Para el modelo básico de remplazo por edad, deduzca la edad óptima de remplazo cuando el costo de mantenimiento $C(u)$ tiene la forma $C(u) = a\sqrt{u}$ para una constante $a > 0$.
23. Suponga que, para el modelo simple de remplazo por edad, el costo de mantenimiento es igual cada año [esto es, $C(u) = a$, para toda $u \geq 0$]. ¿Cuál es la edad óptima de remplazo? ¿Por qué es éste el resultado?
24. El ejército intenta determinar la edad óptima de remplazo de un equipo de campaña. Remplazar ese equipo cuesta \$280 000. El fabricante lo suministra con una rebaja contra la siguiente compra que disminuye el 20% anual. Se estima que los costos de mantenimiento durante el primer año son \$1 000, y que aumentan aproximadamente 18% cada año. Estime la cantidad de años que el ejército debe conservar el equipo antes de remplazarlo.
25. Trate de determinar la edad óptima de remplazo cuando $C(u) = au$ y $S(u) = K - bu$. ¿Qué dificultades encuentra?

12.7 REMPLAZO PLANEADO BAJO INCERTIDUMBRE

El propósito del mantenimiento preventivo es disminuir la probabilidad de que se necesite remplazar un artículo por haber fallado. En la base de esa política está el supuesto que cuesta más reparar o remplazar en el momento de la falla que en un momento predeterminado. Por ejemplo, si la falla se traduce en que se debe detener la línea de producción para determinar la causa de la falla y reparar el problema, mientras que el mantenimiento preventivo se puede hacer en un momento conveniente, cuando el sistema no esté trabajando, el costo de los remplazos planeados será menor que el de los no planeados.

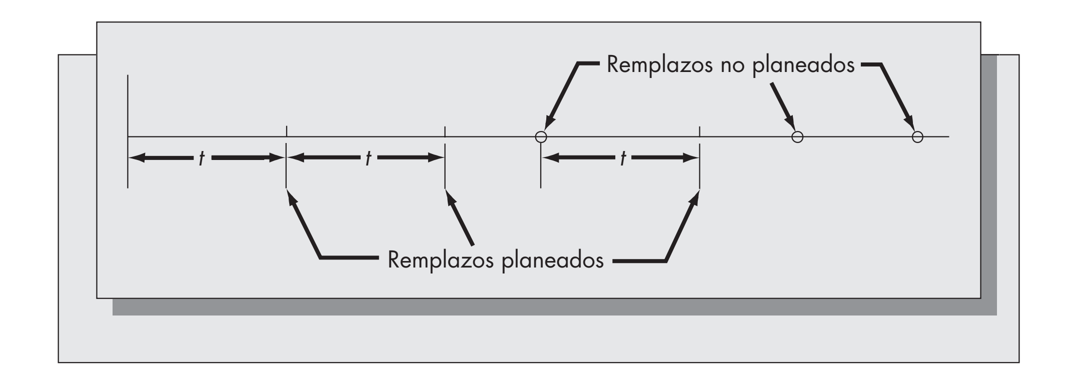
Debido a la propiedad de amnesia de la distribución exponencial, si un artículo o grupo de artículos obedece a una ley exponencial de fallas, no hay ventaja alguna en remplazarlos antes de la falla. En el caso exponencial, la probabilidad de que la falla se presente en un tiempo Δt es igual inmediatamente después de un remplazo planeado que si el artículo ha estado trabajando un intervalo arbitrario de tiempo. Por consiguiente, las estrategias de remplazo planeado sólo pueden tener valor si los artículos muestran envejecimiento, esto es, tienen una función de tasa de fallas en aumento.

Remplazo planeado para un solo artículo

Imaginemos un solo equipo que trabaja continuamente, y cuya vida es una variable aleatoria T con función de distribución acumulada $F(t)$. Suponemos que T es una variable aleatoria continua. También suponemos que cuesta c_1 remplazar el equipo cuando falla, y que cuesta $c_2 < c_1$ remplazarlo antes de que falle. Suponemos que se hacen remplazos planeados exactamente cada t unidades de tiempo después del último remplazo. La meta es determinar el valor óptimo de t que minimice el costo promedio por unidad de tiempo, tanto de remplazos planeados como no planeados.

Un ciclo es el tiempo entre remplazos sucesivos. Como el proceso se "reinicia" después de cada remplazo, independientemente de si fue planeado o no, podemos usar el método de renovación para obtener una ecuación del costo esperado por unidad de tiempo. Esto es,

FIGURA 12-14
Ciclos sucesivos para el remplazo planeado de un solo artículo.



$$E(\text{costo por unidad de tiempo}) = \frac{E(\text{costo por ciclo})}{E(\text{longitud de un ciclo})}.$$

Ya hemos usado argumentos de renovación en este libro (sección 11.6). Este método también se usó en diversos lugares, incluyendo la sección anterior sobre remplazo por edad y en gran parte de los capítulos 4 y 5 sobre modelado de inventarios. Los ciclos sucesivos de remplazo se ilustran en la figura 12-14.

Tenemos que,

$$E(\text{costo por ciclo}) = c_1P\{\text{remplazo debido a falla}\} + c_2P\{\text{remplazo planeado}\}.$$

Observamos que $P\{\text{remplazo debido a falla}\} = P\{T \le t\} = F(t)$, y que $P\{\text{remplazo planeado}\} = P\{T > t\} = 1 - F(t)$, donde T es la vida del artículo puesto en servicio al final del ciclo anterior. Por consiguiente,

$$E(\text{costo por ciclo}) = c_1F(t) + c_2[1 - F(t)].$$

Sea T el tiempo de falla del artículo que se puso en servicio al final del ciclo anterior. Es claro entonces que el siguiente remplazo se tendrá en el tiempo mín(T, t). Por consiguiente,

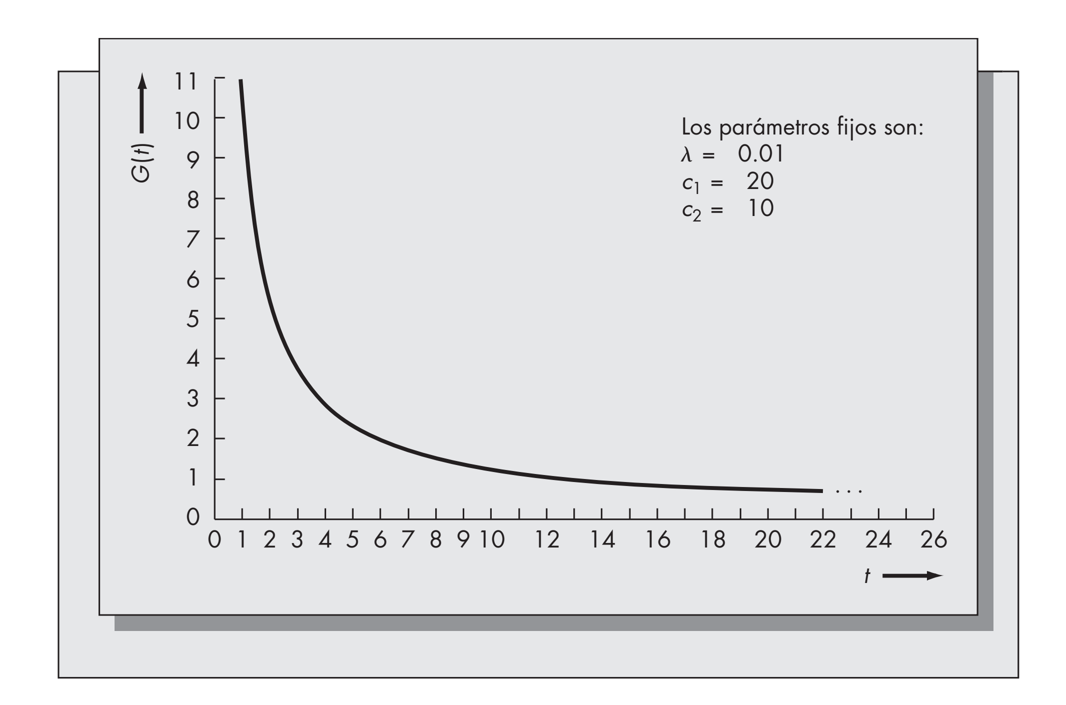
$$\begin{aligned} E(\text{longitud de ciclo}) = E[\text{mín}(T, t)] &= \int_0^\infty \text{mín}(x, t)f(x)\,dx \\ &= \int_0^t xf(x)\,dx + t\int_0^\infty f(x)\,dx \\ &= \int_0^t xf(x)\,dx + t[1 - F(t)]. \end{aligned}$$

En consecuencia, el costo esperado por unidad de tiempo, digamos $G(t)$, se expresa como sigue:

$$G(t) = \frac{c_1F(t) + c_2[1 - F(t)]}{\int_0^t xf(x)dx + t[1 - F(t)]}.$$

El objetivo es determinar t que minimice a $G(t)$. La optimización puede ser tediosa, según la forma de la distribución $F(t)$, del tiempo de vida.

FIGURA 12-15
La función $G(t)$.



Ahora demostraremos que no se obtiene ninguna ventaja con el remplazo planeado cuando la distribución del tiempo de vida es exponencial. Supongamos que $F(t) = 1 - e^{-\lambda t}$. Entonces, la longitud esperada de cada ciclo es

$$\int_0^t xe^{-\lambda x}\,dx + te^{-\lambda t} = \frac{1}{\lambda}[1 - e^{-\lambda t}(1 + \lambda t)] + te^{-\lambda t}$$

$$= \frac{1}{\lambda}(-e^{-\lambda t}).$$

(La expresión para $\int_0^t xe^{-\lambda x}\,dx$ puede obtenerse integrando por partes o puede encontrarse en una tabla de integrales.) Por consiguiente,

$$G(t) = \frac{c_1(1 - e^{-\lambda t}) + c_2\, e^{-\lambda t}}{\frac{1}{\lambda}(1 - e^{-\lambda t})} = \frac{c_1 - (c_1 - c_2)\, e^{-\lambda t}}{\frac{1}{\lambda} - \frac{1}{\lambda}e^{-\lambda t}}.$$

Cuando $t \to \infty$, el término $e^{-\lambda t} \to 0$, así que $G(\infty) = \lambda c_1$. Además, al sustituir $t = 0$ se obtiene $G(0) = \infty$. Es posible demostrar, por cálculo infinitesimal o por cálculo directo, que la función $G(t)$ es monótonamente decreciente; un caso característico se aprecia en la figura 12-15. Por consiguiente, la solución óptima es $t = \infty$, que significa que nunca debe hacerse un remplazo planeado.

Hemos demostrado que si la distribución de la vida es exponencial (tasa constante de fallas), no se consiguen economías al remplazar un artículo antes del momento en que falle. Esto también es válido si la tasa de fallas es decreciente.

Ejemplo *12.11* Harley Brown, Inc., una gran empresa camionera, mantiene registros detallados sobre la mortalidad de los neumáticos que usan sus camiones. Un análisis estadístico de los datos muestra que la vida de un neumático, medida en miles de millas de empleo, se aproxima muy bien con la ley de probabilidades de Weibull con parámetros $\alpha = .00235$ y $\beta = 2.3$. La empresa estima un costo de $450 si un neumático

TABLA 12-2 Probabilidades de falla para el ejemplo 12.11

Tiempo de vida (miles de millas)	Probabilidad de falla	Tiempo de vida (miles de millas)	Probabilidad de falla
1	.0023	14	.0625
2	.0092	15	.0580
3	.0175	16	.0527
4	.0264	17	.0469
5	.0354	18	.0408
6	.0440	19	.0348
7	.0517	20	.0291
8	.0582	21	.0238
9	.0632	22	.0191
10	.0664	23	.0150
11	.0679	24	.0116
12	.0677	25	.0087
13	.0658		

falla cuando se está usando. Esto se debe al tiempo perdido y a la responsabilidad potencial en un accidente. Los neumáticos remplazados antes de su falla cuestan 220 dólares cada uno. La empresa quisiera determinar el momento óptimo para remplazar los neumáticos.

La dificultad en el cálculo de la solución óptima para este problema estriba en determinar una ecuación para el término $\int_0^t x f(x)\, dx$ que aparece en el denominador de $G(t)$.

Para solventar el problema de encontrar una expresión analítica para esta integral, obtendremos una aproximación discreta a la ley de fallas y efectuaremos el cálculo como si la distribución de la vida fuera discreta y no continua. La probabilidad de falla dentro de las primeras 1 000 millas es

$$P\{T \leq 1\} = 1 - F(1) = .0023.$$

La probabilidad de falla después de 1 000 millas, pero antes de 2 000 millas de uso, es

$$P\{T \leq 2\} - P\{T \leq 1\} = R(1) - R(2) = .0092.$$

El resto de las probabilidades de falla se calculan en forma parecida y se ven en la tabla 12-2.

Usando estas probabilidades discretas ya podemos calcular $G(t)$ en forma directa. El término de la esperanza parcial

$$\int_0^t x f(x)\, dx \approx \sum_{k=1}^{t} k p_k,$$

donde las probabilidades p_k son las de la tabla 12-2. Obsérvese que en esta aproximación se supone que todas las fallas se presentan en múltiplos de 1 000 millas.

El resto de los términos que forman $G(t)$ se puede obtener directamente a partir de la función de confiabilidad de Weibull,

$$R(t) = e^{-\alpha t^{\beta}}.$$

Por consiguiente, la forma aproximada de $G(t)$ se puede escribir como sigue:

$$G(t) = \frac{c_1 - (c_1 - c_2)R(t)}{\sum_{k=1}^{t} k p_k + tR(t)}.$$

$G(t)$ aparece en la tabla 12-3 para los valores de los parámetros $\alpha = .00235$, $\beta = 2.3$, $c_1 = 450$ y $c_2 = 220$ dólares. Parece que la función es convexa y tiene su mínimo cuando $t = 13$. Por consiguiente, la política óptima es remplazar los neumáticos después de unas 13 000 millas de uso. El valor de la función objetivo en la solución óptima es $G(13) = 33.21$. Esto significa que en la solución óptima el costo de remplazo es \$33.21 por mil millas de uso por neumático.

TABLA 12-3 **La función $G(t)$ del ejemplo 12.11**

t	$G(t)$	t	$G(t)$
1	220.54	14	33.24
2	111.45	15	33.35
3	75.91	16	33.53
4	58.82	17	33.72
5	49.14	18	33.93
6	43.20	19	34.13
7	39.38	20	34.32
8	36.89	21	34.48
9	35.27	22	34.62
10	34.25	23	34.74
11	33.64	24	34.84
12	33.33	25	34.91
13	33.21		

Remplazo en bloque para un grupo de artículos

En ciertas circunstancias es más económico remplazar grupos de artículos a la vez, y no uno por uno. En el ejemplo anterior se demostró que la política óptima fue remplazar un neumático de camión después de aproximadamente 13 000 millas de uso. Dependiendo del tiempo y los gastos que intervienen en el cambio de neumáticos, podría ser más económico remplazar todos los neumáticos de un camión cuando se hace un cambio. Los costos de transportar el camión a un taller de servicio, colocarlo en un gato hidráulico y pagar a un técnico para que monte y haga el balanceo de los neumáticos podría ser comparable al costo del neumático mismo. Si se remplazan todos los neumáticos a la vez, se incurriría con menos frecuencia en este costo que si se remplazaran uno por uno.

En esta sección consideraremos un modelo para determinar el tiempo óptimo para remplazar un grupo completo de artículos. Con el propósito de evitar las matemáticas intrincadas que salen del ámbito de este libro, supondremos que la vida de cada unidad en funcionamiento es una variable aleatoria discreta con una distribución conocida. Esto es, supondremos que p_k es la probabilidad de que un artículo falle en el periodo k, suponiendo que se puso en servicio en el periodo 0. Esas probabilidades se pueden estimar en forma directa a partir de datos históricos, o calcular a partir de una distribución continua, como en el ejemplo 12.11.

Suponga que se ponen en servicio n_0 artículos en el tiempo 0. Suponga que no hay remplazo en bloque, y que todos los artículos que fallan en un periodo se remplazan al final de ese periodo. También supondremos, por simplicidad, que p_k es la proporción real de unidades con k periodos de edad que fallan. Entonces, la cantidad de fallas que ocurren en el periodo 1 es $n_1 = n_0 p_1$.

En el periodo 2 la proporción del grupo original de artículos que fallan es $n_0 p_2$ y la de artículos que se ponen en servicio en el periodo y que fallan es $n_1 p_1$. Por consiguiente, la cantidad esperada de fallas en el periodo 2 es $n_2 = n_0 p_2 + n_1 p_1$. Si continuamos con este argumento llegaremos a

$$n_k = n_0 p_k + n_1 p_{k-1} + \cdots + n_{k-1} p_1.$$

Ahora suponga que los remplazos individuales cuestan a_1 cada uno, y que todo el bloque de n_0 se puede remplazar con un costo a_2. Si todos los artículos n_0 se remplazaran al final de cada periodo, el costo de cada periodo sería $a_2 + a_1 n_1$. Si todos los n_0 artículos se remplazaran al final de cada dos periodos, el costo incurrido cada dos periodos sería $a_2 + a_1(n_1 + n_2)$, es decir, un costo promedio por periodo igual a $[a_2 + a_1(n_1 + n_2)]/2$. De igual manera, el costo promedio por periodo de remplazar los n_0 artículos al cabo de k periodos es

$$G(k) = \frac{a_2 + a_1 \sum_{j=1}^{k} n_j}{k}.$$

La cantidad óptima de periodos para remplazar los n_0 artículos es el valor de k que minimiza a $G(k)$. El valor mínimo de $G(k)$ debe compararse con el costo esperado por periodo, suponiendo que todos los artículos se remplazan cuando fallan. Sea

$$E(T) = \sum_{k=1}^{\infty} kp_k$$

la vida esperada de un solo artículo. Entonces, $\lambda = 1/T$ es la tasa de fallas de un solo artículo. Por consiguiente, el costo de remplazar artículos uno por uno es $a_1\lambda$ por cada artículo, es decir $n_0a_1\lambda$ para el conjunto de artículos. Esto debe compararse con el valor óptimo de $G(k)$ para determinar si la estrategia de remplazo en bloque es económica.

Ejemplo *12.12* Un gran letrero luminoso está formado por 8 000 focos. Cuesta 2 dólares remplazar cada foco en cuanto falla, pero podría ser de .30 de dólar por pieza si se remplazaran todos a la vez. De acuerdo con la experiencia, los focos fallan siguiendo la ley de probabilidades de la siguiente tabla:

Meses de servicio	**Probabilidad de falla**
1	.02
2	.03
3	.03
4	.05
5	.08
6	.09
7	.07
8	.10
9	.11
10	.13
11	.15
12	.14

El primer paso es calcular n_k. Tenemos que

$$n_0 = 8\,000,$$
$$n_1 = n_0p_1 = (8\,000)(.02) = 160,$$
$$n_2 = n_0p_2 + n_1p_1 = (8\,000)(.03) = (160)(.02) = 243,$$

y así sucesivamente.

Estos valores se usan para calcular $G(k)$. En la tabla 12-4 aparecen los resultados de los cálculos. Si se usan los valores de $a_2 = (0.30)(8\,000) = \$2\,400$ y $a_1 = 2$, se ve que el momento óptimo para remplazar el conjunto de los focos sería pasados 4 meses, con un costo mensual esperado igual a 1 135 dólares.

Es interesante comparar el remplazo en bloque con una política de remplazo de los focos en el momento que fallan. Cada foco tiene una vida esperada que se expresa como sigue:

$$E(T) = \sum_{k=1}^{12} kp_k = 8.22 \text{ meses.}$$

Por lo tanto, la tasa de fallas es 1/8.22 = 0.12165 fallas por mes por foco. Para 8 000 focos esto equivale a una cantidad promedio de fallas igual a 973.24 por mes. El costo resultante de remplazo es \$1 946.47 por mes, bastante más que remplazar los focos en bloque cada 4 meses, cuyo costo mensual esperado es \$1 135.

TABLA 12-4 ***G*(*k*) para el ejemplo 12.12**

k	P_k	n_k	G_k
1	.02	160.0	2 720.0
2	.03	243.2	1 603.2
3	.03	249.2	1 235.2
4	.05	417.1	1 135.0
5	.08	671.1	1 176.4
6	.09	778.4	1 239.8
7	.07	654.6	1 249.7
8	.10	930.5	1 326.1
9	.11	1 064.0	1 415.2
10	.13	1 298.4	1 533.4
11	.15	1 542.9	1 674.5
12	.14	1 562.4	1 795.4

Problemas para la sección 12.7

26. Vea el ejemplo 12.11 de Harley Brown, Inc. Sin hacer los cálculos, comente cuál sería el efecto probable sobre la política óptima de remplazo si se cambiaran los parámetros a $c_1 = \$800$, $c_2 = \$300$ y $\beta = 1$.

27. Repita los cálculos de la empresa Harley Brown, Inc., con los siguientes valores de parámetros: $\alpha = .0516$, $\beta = 1.8$, $c_1 = 1\,000$ y $c_2 = 600$.

28. Una pieza de equipo costosa se emplea en la operación de enmascarado durante la fabricación de semiconductores. Un capacitor del equipo falla al azar. Ese capacitor cuesta \$7.50, pero si se quema cuando la máquina está funcionando, el proceso de producción debe detenerse. En este caso, se estima que el costo de remplazo es \$150. De acuerdo con la experiencia, se estima que la distribución de la vida del capacitor es:

Número de meses de servicio	Probabilidad de falla
1	.08
2	.12
3	.16
4	.26
5	.22
6	.16

¿Con qué frecuencia deben remplazarse los capacitores para minimizar el costo mensual esperado de remplazo planeado y no planeado?

29. Un gran órgano electrónico contiene 100 fusibles. Debido al consumo de corriente, los fusibles se queman con una tasa bastante regular, pero los fusibles nuevos duran más que los anteriores. La distribución de probabilidades del tiempo de vida de un fusible se aproxima bastante bien a la ley de Weibull con $\alpha = 0.0204$ y $\beta = 1.8$. Suponga que t es el tiempo de funcionamiento en horas. Los fusibles cuestan \$1.35 cada uno cuando se remplazan en bloque, y \$12 cada uno cuando se remplazan justo después de fallar.

    a) Exprese la distribución del tiempo de vida en forma de distribución discreta, suponiendo que t se mide en horas. (Siga el procedimiento usado en el ejemplo 12.11 para la Harley Brown, Inc.)

b) Determine el tiempo óptimo para remplazar los 100 fusibles, y el costo promedio por hora de esa política.

c) Compare la respuesta obtenida en el inciso *b*) con el costo de remplazar los fusibles en el momento en que fallan. ¿Cuál política recomienda usted?

30. Los neumáticos que fallan en servicio tienen costos de remplazo bastante mayores que los que se cambian antes de que fallen. Para un camión con 18 neumáticos, la falla en carretera cuesta $300, mientras que si se cambian los 18 neumáticos antes de que fallen, el costo es $75 cada uno. La probabilidad de falla es la siguiente:

Número de millas	Probabilidad de falla
0-5 000	.05
5 001-10 000	.15
10 001-15 000	.20
15 001-20 000	.40
20 001-25 000	.20

Si los neumáticos se remplazan sólo en múltiplos de 5 000 millas, ¿cuál es la política óptima de remplazo por edad?

31. La Marina usa cierto tipo de tubo al vacío en un dispositivo de sonar. De acuerdo con la experiencia, esos tubos tienen el siguiente comportamiento de fallas:

Número de meses de operación	Probabilidad de falla
1	.1
2	.1
3	.2
4	.1
5	.3
6	.2

Si el tubo falla durante el funcionamiento, cuesta $200 cada uno, pero si se puede remplazar antes de que falle, cuesta $50. Calcule la estrategia óptima de remplazo.

32. Un periódico local se imprime en una máquina con cinta, que se puede romper o acabársele la tinta durante el funcionamiento. Las cintas pueden remplazarse por $7.50, pero si fallan cuando se imprime el periódico, el costo se estima en $25, por la demora en la publicación. La distribución de fallas de las cintas es la siguiente:

Semanas de uso	Probabilidad de falla
1	.1
2	.2
3	.3
4	.4

Determine el tiempo óptimo para remplazar las cintas.

33. Mactronics produce robots industriales. Cada robot contiene una parte cuya vida es 5 años, como máximo, y un año como mínimo. Las vidas de 1 a 5 años tienen igual probabilidad.

Si la parte falla durante la operación, los costos de remplazo se estiman en $400, pero la parte se puede remplazar antes de que falle por $50. ¿Cuándo hay que remplazar la parte? [Sugerencia: la distribución de vida es uniforme en {1, 2, 3, 4, 5}. Suponiendo que las variables sean discretas, esto indica que $f(x) = \frac{1}{5}$ para $x = 1, 2, \ldots, 5$.]

34. Una empresa compró 30 robots a Mactronics del problema anterior. Si se remplazan en bloque, las partes cuestan $20 cada una, pero cuestan $400 cada una cuando se remplazan después de fallar.

a) Determine la estrategia óptima de remplazo en bloque.

b) Compárela con el costo de remplazar los artículos cuando fallan.

c) Compare el costo de la política de remplazo en bloque que se obtuvo en el inciso *a*) con la solución que se obtuvo en el inciso *b*). ¿Es preferible la estrategia de remplazo en bloque frente a una de remplazo individual?

*12.8 ANÁLISIS DE POLÍTICAS DE GARANTÍA

Un tema importante que se relaciona con la confiabilidad del equipo en funcionamiento es la protección que se da al consumidor, quien sufre las fallas del equipo antes de su tiempo de vida calculada. Los compradores y los vendedores perciben las garantías en forma distinta. Desde el punto de vista del vendedor, la garantía es un medio de limitar su responsabilidad, al especificarse las responsabilidades del consumidor. Entre esas responsabilidades se encuentran el uso adecuado del producto y el apego a las advertencias. Desde la perspectiva de la mercadotecnia, la garantía también puede servir como inductor a la compra del producto. Desde el punto de vista del comprador, la garantía es un medio de reducir o eliminar la penalización económica cuando el producto no puede funcionar bien dentro de un tiempo razonable. Las garantías tienen especial importancia para el consumidor de productos que probablemente manifiesten grandes tasas de falla, al principio de su vida.

En esta sección se presentan modelos matemáticos para determinar el valor económico de una garantía. Esos modelos se podrían usar para calcular la porción del costo de un artículo que se puede atribuir razonablemente a los costos de satisfacer un compromiso de garantía. En los modelos consideraremos tanto la estructura de la garantía como la confiabilidad del producto, para calcular el valor de la garantía.

Debe distinguirse entre artículos reparables y no reparables. Entre los no reparables se incluyen la mayor parte de los componentes electrónicos, artículos en los que la falla equivale a su propia destrucción (fundido de un foco o voladura de un neumático, por ejemplo), o los artículos que no suelen repararse, sino remplazarse (como los acumuladores que se descargan). La mayor parte de los electrodomésticos, como máquinas lavadoras, televisores, etcétera, están en la categoría de artículos reparables. Los artículos reparables que fallan durante el periodo de garantía se reparan y no se remplazan. En los modelos matemáticos que se presentan en esta sección se supone que los artículos no son reparables.

Las garantías para bienes no reparables de consumo toman por lo general dos formas. Una es de remplazo sin costo: si sucede una falla durante el periodo de garantía, se suministra sin costo un artículo nuevo. El segundo tipo de garantía es la de prorrata. En este caso, al consumidor se le da una compensación proporcional al tiempo que resta del periodo de garantía. La compensación se usa para reducir el costo de un artículo remplazable.

La garantía de reposición sin costo

Supóngase que una pieza de un equipo en funcionamiento se pone en servicio y falla en forma totalmente aleatoria (esto es, apegándose a una ley exponencial de fallas), con una frecuencia

conocida de falla λ. Se supone que el artículo trabaja en forma continua. Si el equipo trabaja en forma intermitente, el tiempo se debe expresar en horas de funcionamiento y no en horas transcurridas.

Usaremos la siguiente notación:

T = Tiempo de vida de un artículo elegido al azar.

λ = Tasa de fallas de un artículo elegido al azar.

$F(t)$ = Función de distribución acumulada de la variable aleatoria t.

C_1 = Costo de compra de un artículo nuevo con garantía de remplazo sin costo.

K = Costo de compra de un artículo nuevo sin garantía alguna.

W_1 = Tiempo de vigencia de la garantía de reposición sin costo, después de la compra.

Si ocurre una falla durante el periodo de garantía, la parte se remplaza sin cargo. Supóngase que el consumidor compra un artículo nuevo cuando sucede una falla después de haber expirado la garantía. El artículo nuevo tiene una garantía idéntica de remplazo sin costo. Sea Y una variable aleatoria que representa el tiempo entre compras sucesivas. En la figura 12-6 se ve que

$$Y = W_1 + \text{Tiempo hasta la primera falla, después de haber expirado la garantía.}$$

Es posible demostrar que

$$E(Y) = W_1 + 1/\lambda.$$

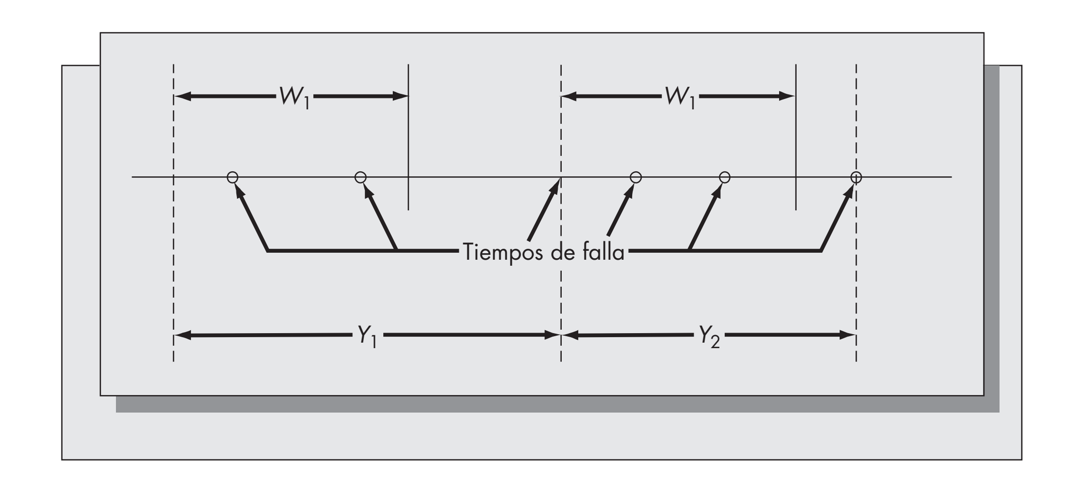
Esta ecuación de $E(Y)$ es válida cuando la distribución de la vida es exponencial. Se debe a la propiedad de la ley exponencial de fallas que el tiempo a la primera falla después resulta un tiempo fijo (llamado tiempo de recurrencia directa en la teoría de probabilidad) con la misma distribución que el tiempo entre dos fallas sucesivas. Este resultado *sólo* es cierto si la ley de fallas es exponencial.

Ahora calcularemos el costo por unidad de tiempo para un artículo que se remplaza una cantidad infinita de veces, y que tiene una garantía de remplazo sin costo.

Cada vez que se compra un artículo después de que expira la garantía, se define el inicio de un nuevo ciclo. Como el costo por ciclo es C_1, en consecuencia el costo promedio por unidad de tiempo con la garantía de remplazo sin costo es

$$\frac{C_1}{E(Y)} = \frac{C_1}{W_1 + 1/\lambda} = \frac{\lambda C_1}{\lambda W_1 + 1}.$$

FIGURA 12-16
Ciclos de remplazo para la garantía de remplazo sin costo.



Sin la garantía, el costo de remplazo por unidad de tiempo es simplemente λK. Sea C_1^* el costo de un artículo con garantía de remplazo sin costo que es indiferente al costo de un artículo sin la garantía, K. En ese caso, C_1^* es la solución de

$$\lambda K = \frac{\lambda C_1^*}{\lambda W_1 + 1}$$

esto es

$$C_1^* = (\lambda W_1 + 1)K.$$

Por definición, $C_1^* - K$ es el valor económico de la garantía. Si $C_1 < C_1^*$, se debe comprar la garantía.

La garantía de prorrata

Adoptaremos toda la notación presentada con anterioridad en esta sección. También definiremos

C_2 = Costo de un artículo nuevo con garantía de prorrata.

W_2 = Periodo efectivo de garantía con una garantía de prorrata.

Consideremos un artículo que se compra con garantía de prorrata que falla en un tiempo aleatorio T. Existen dos casos:

Caso 1: $T < W_2$. En este caso, la fracción del periodo de garantía que ya expiró es T/W_2. El costo de remplazar el artículo es, entonces, $C_2(T/W_2)$.

Caso 2: $T \geq W_2$. Aquí, la garantía expiró y el costo de remplazo es C_2.

Ambos casos pueden representarse matemáticamente con la expresión

$$\frac{C_2}{W_2}\,\text{mín}(W_2, T).$$

Para determinar el costo en el ciclo esperado de vida, necesitamos encontrar una ecuación para $E[\text{mín}(W_2, T)]$. Tenemos que

$$\begin{aligned} E[\text{mín}(W_2, T)] &= \int_0^\infty \text{mín}(W_2, t)\lambda e^{-\lambda t}\,dt \\ &= \int_0^{W_2} t\lambda e^{-\lambda t}\,dt + W_2\int_{W_2}^\infty \lambda e^{-\lambda t}\,dt, \end{aligned}$$

donde $\lambda e^{-\lambda t}$ es la densidad de probabilidad del tiempo hasta la falla, T.

La primera integral requiere integración por partes, y la segunda es la función de confiabilidad evaluada en W_2. Es fácil demostrar que

$$\int_0^{W_2} t\lambda e^{-\lambda t}\,dt = \frac{1}{\lambda}[1 - e^{-\lambda W_2}(1 + \lambda W_2)]$$

y

$$W_2\int_{W_2}^\infty \lambda e^{-\lambda t}\,dt = W_2 e^{-\lambda W_2}.$$

Al combinar los términos se ve que

$$E[\text{mín}(W_2, T)] = \frac{1}{\lambda}(1 - e^{-\lambda W_2}).$$

La garantía de prorrata comienza de nuevo con cada compra. Por consiguiente, cada compra comienza un nuevo ciclo. Así, el costo esperado por unidad de tiempo después de obtener una garantía de prorrata es

$$\frac{C_2}{W_2}\frac{E[\text{mín}(W_2, T)]}{E(T)} = \frac{C_2}{W_2}\frac{(1/\lambda)(1- e^{-\lambda W_2})}{1/\lambda} = \frac{C_2(1 - e^{-\lambda W_2})}{W_2}.$$

El costo de la garantía de prorrata que es indiferente al costo del artículo sin garantía, C_2^*, es la solución de

$$\frac{C_2^*(1 - e^{-\lambda W_2})}{W_2} = \lambda K,$$

o bien

$$C_2^* = \frac{\lambda K W_2}{1 - e^{-\lambda W_2}}.$$

El valor de la garantía de prorrata es $C_2^* - K$. Si $C_2 > C_2^*$, no debe comprarse la garantía de prorrata, y si $C_2 < C_2^*$, dicha compra sí debe realizarse. Con esto se supone que el consumidor está de acuerdo en basar su decisión en valores esperados. Si al consumidor no le gusta tomar riesgos, el valor de indiferencia será un poco mayor que C_2^*.

Ejemplo *12.13* Usted está considerando comprar un acumulador para su coche. Encuentra el mismo acumulador en tres tiendas distintas. En la tienda A se vende en $21, y no ofrece garantías. En la tienda B se vende en $40, y ofrece remplazo gratis si se descarga dentro de los dos primeros años de funcionamiento. En la tienda C se vende también a $40, pero ofrece una garantía de prorrata por la duración anticipada del acumulador, que se anuncia como de 5 años. La tasa de fallas del acumulador depende del uso y las condiciones, pero usted calcula, según su experiencia, que el tiempo entre fallas es más o menos una vez cada 3 años.

Determinaremos los valores de C_1^* y C_2^* para ambas garantías. En el caso de la garantía de remplazo total, tenemos que

$$K = 21,$$
$$C_1 = 40,$$
$$\lambda = \tfrac{1}{3},$$
$$W_1 = 2.$$

De aquí, se obtiene

$$C_1^* = (\lambda W_1 + 1)K = [(\tfrac{1}{3})(2) + 1](21) = \$35.$$

Esto indica que el valor de la garantía de remplazo sin costo es $35 − $21 = $14, que es menor que la diferencia de precios de $19. Con base en los costos esperados, se debe preferir el acumulador sin garantía (tienda A) frente al que tiene garantía de remplazo sin costo (tienda B).

Para el caso de la garantía de prorrata que ofrece la tienda C, $W_2 = 5$, y los parámetros restantes, son los antes definidos. En este caso, el resultado es

$$C_2^* = \frac{\lambda K W_2}{1 - e^{-\lambda W_2}} = \frac{\frac{1}{3}(21)(5)}{1 - e^{-5/3}} = \$43.15.$$

El valor de la garantía de prorrata es $43.15 − $21.00 = $22.15 dólares, mayor que la diferencia entre el precio del acumulador con y sin la garantía, y es mayor que el valor de la garantía con remplazo sin costo. Por consiguiente, con base en este análisis, la garantía de prorrata es la opción preferida.

Extensiones y críticas

Una crítica al ejemplo 12.13 es por el supuesto de que la ley de fallas de los acumuladores es exponencial. Esto significa que un acumulador nuevo tiene la misma probabilidad de fallar en su primer año de funcionamiento, que la que tiene uno con 4 años de edad de fallar en su quinto año de funcionamiento. En el caso de los acumuladores, parece que sería más exacta una ley de fallas que tuviera en cuenta el envejecimiento, como la de Weibull. Los modelos descritos pueden ampliarse para abarcar tipos más generales de distribuciones de fallas, pero los cálculos que se requieren para determinar una política óptima son muy complejos. En particular, debe determinarse la función de renovación, que por lo general es un cálculo difícil. En el problema 41 se presenta un ejemplo para el caso en el que la ley de fallas sigue a una distribución de Erlang.

Los dos tipos de garantías descritos en esta sección, la de remplazo sin costo y la de prorrata, son los que se usan en la mayor parte de las ventas al consumidor de artículos no reparables. En el contexto militar se acostumbra mucho otro tipo de garantía, llamado *garantía de mejoramiento de confiabilidad* (RIW, por sus siglas en inglés). En este caso, el proveedor acuerda reparar o remplazar artículos que fallen dentro de un periodo especificado de garantía, y suministra uno o varios conjuntos de piezas de repuesto, y quizá una o más instalaciones de reparación. Este tipo de garantía tiene por objeto dar un incentivo para que el proveedor otorgue una confiabilidad inicial alta, y para que, si es posible, mejore la confiabilidad de unidades ya existentes.

Para los artículos reparables, las garantías cubren por lo general todo el costo de las reparaciones, o parte de éste, durante el periodo garantizado. El análisis de las garantías para artículos reparables es más complejo que para los no reparables. Son posibles distintos niveles de reparación, y la probabilidad de falla durante el periodo de garantía puede sujetarse al uso, que puede variar mucho de un usuario a otro.

Asimismo, para los artículos reparables, un aspecto estrechamente relacionado con las garantías es el de los contratos de servicio. La diferencia principal entre una garantía y un contrato de servicio es que, generalmente, el costo de la garantía está incluido en el precio de compra del artículo, mientras que un contrato de servicio es un artículo adicional, cuya adquisición es una opción para el comprador. Se puede concebir un contrato de servicio como una garantía ampliada más allá del periodo garantizado normal. El tratamiento analítico que se describe aquí también puede aplicarse a la determinación del valor económico de un contrato de servicio. Sin embargo, deben tenerse en cuenta leyes complejas de falla, y los diversos costos asociados con distintos niveles de reparación. Además, el precio de los contratos de servicio al cliente depende mucho del supuesto que afirma que el cliente es adverso al riesgo. Esto es, un consumidor normal prefiere pagar más de lo que vale un contrato de servicio, en el sentido de valor esperado, con tal de reducir el riesgo de incurrir en una reparación costosa. En ese sentido, el contrato de servicio desempeña el papel de una póliza de seguro. Quizá sea la tendencia de los consumidores a evitar riesgos lo que justifica los precios tan altos que se cobran en los contratos de servicio para algunos productos, por ejemplo, las grabadoras de videocintas. (Véase Lund, 1978, pp. 4-63.)

Problemas para la sección 12.8

35. Para el ejemplo 12.13, ¿qué valor del periodo de garantía iguala a la garantía de remplazo total con la alternativa de no usar garantía? Esto es, ¿para qué valor de W_1, al consumidor le es indistinto comprar en la tienda A que en la tienda B?

36. Para el ejemplo 12.13, ¿qué valor del periodo de garantía iguala la garantía de remplazo total con la de remplazo de prorrata? Esto es, ¿para qué valor de W_1 al cliente le es indistinto comprar en la tienda B o en la C, suponiendo que $W_2 = 5$?

37. Un productor de calculadoras de bolsillo estima que éstas fallan con una tasa de una cada 5 años. Vende a $25 cada calculadora, con una garantía de remplazo sin costo, pero se les puede comprar de otro proveedor a $18.50 sin garantía. ¿Vale la pena comprar la calculadora con garantía?
38. Para el problema 37, ¿qué duración de periodo de garantía iguala los costos de remplazo de la calculadora con y sin la garantía?
39. Zemansky's vende neumáticos con garantía de prorrata. Garantiza que los neumáticos duran 50 000 millas con devolución basada en las lonas restantes del neumático. En promedio, los neumáticos fallan a las 35 000 millas de uso. Suponga que se venden a $50 cada uno con la garantía. Si las fallas se presentan totalmente al azar, ¿cuál sería un precio consistente de los neumáticos si se vendieran sin garantía?
40. Habard's, una cadena de ferreterías, vende diversas herramientas y artículos para el hogar. Una de sus mejores llaves se vende a $5.50. Por $1.50 más, Habard's incluye una garantía de remplazo sin costo durante 3 años. Las llaves son para uso rudo y, con base en la experiencia, fallan al azar, con una frecuencia de una vez cada 8 años. ¿Vale la pena comprar la garantía?
41. Considere el caso en el que el mecanismo de falla del producto no se apega a la ley exponencial. En ese caso, el costo con la garantía de remplazo gratuita, que es indiferente a la compra del artículo sin garantía, está dado por:

$$C_1^* = K[M(W_1) + 1],$$

donde $M(t)$ se llama función de renovación.

Si el tiempo entre fallas, T, sigue una ley de Erlang con parámetros λ y 2, entonces

$$M(t) = \frac{\lambda t}{2} - 0.25 + 0.25e^{-2\lambda t} \qquad \text{para toda } t \geq 0.$$

(Véase, por ejemplo, Barlow y Proschan, 1965, p. 57.)

a) Para el ejemplo 12.13 de esta sección, determine el valor de indiferencia del artículo, con una garantía de reposición sin costo, cuando la ley de fallas tiene una distribución de Erlang. Suponga que $\lambda = \frac{2}{3}$ para llegar al mismo valor de $E(T)$ que en el ejemplo.

b) ¿Es mayor o menor el valor de la garantía que el caso exponencial correspondiente? Explique en forma intuitiva el resultado.

12.9 CONFIABILIDAD DEL SOFTWARE

La confiabilidad del software es un problema cuyas características son distintas a las de los problemas con el hardware. Normalmente, los programas nuevos tienen algunos errores. En el caso ideal, se querría eliminar todos los errores antes de lanzar el programa al mercado, pero esto es imposible. Es más razonable lanzar el programa cuando se haya reducido la cantidad de errores hasta un nivel aceptable. Sin embargo, predecir la cantidad de errores que restan es un problema difícil.

No se puede exagerar la importancia de la confiabilidad de los programas. Citando a *The Wall Street Journal* (Davis, 1987):

> El más pequeño error de un software frecuentemente puede dar al traste con la máquina más poderosa, lo que frecuentemente genera desastrosas consecuencias. Durante los últimos 5 años los defectos en los programas han matado marineros, paralizado pacientes, lastimado empresas y amenazado

con causar el colapso del mercado de bonos gubernamentales. Es probable que esos problemas crezcan a medida que la industria y los militares confíen dada vez más en los programas para hacer trabajar sistemas de complejidad fenomenal, incluyendo el sistema de defensa antimisiles, el de la “Guerra de las Galaxias” propuesto por el presidente Reagan.

Se han propuesto varios modelos para estimar la confiabilidad de los sistemas informáticos. Sin embargo, no los presentaremos aquí con detalle, porque todavía no se determina su utilidad. Jelinski y Moranda (1972) han sugerido el siguiente método. Sea N el contenido inicial total de errores (es decir, la cantidad de errores) en el programa. Al probar el programa, la cantidad de errores se reduce. Suponen que la tasa de fallas (esto es, la probabilidad de localizar un error) es proporcional a la cantidad de errores restantes en el programa, con la

Aplicación instantánea

EL MANTENIMIENTO CENTRADO EN LA CONFIABILIDAD MEJORA LAS OPERACIONES EN LA PLANTA NUCLEAR DE THREE MILE ISLAND

La instalación nuclear de Three Mile Island, ubicada en el río Susquehanna, a unas 10 millas de Harrisburg, Pennsylvania, es notable por un aspecto. Fue el sitio del peor accidente en una planta nucleoeléctrica en Estados Unidos. En marzo de 1979, la Unidad 2 sufrió una fusión del núcleo cuando fallaron los sistemas de seguridad y no levantaron las varillas de combustible del núcleo. La instalación permaneció cerrada los siguientes 6 y medio años como resultado del accidente, y reabrió en octubre de 1985. La central, operada por GPU Nuclear Corporation, ha establecido una de las marcas más impresionantes en la industria desde que reabrió. Según Fox *et al*. (1994), fue clasificada en 1989 como la primera en el mundo de acuerdo con su factor de capacidad (proporción de tiempo en funcionamiento).

En 1987, GPU comenzó a considerar las ventajas de un mantenimiento centrado en la confiabilidad (RCM, por sus siglas en inglés) para su mantenimiento preventivo. Identificaron 28 de un total de 134 sistemas como candidatos viables para este método. Esos 28 sistemas incluían la turbina principal, el sistema de agua de refrigeración, el generador principal, el agua de circulación, y otros más. El proceso de mantenimiento centrado en la confiabilidad se basaba en los cuatro principios básicos siguientes:

- Preservar las funciones del sistema.
- Identificar las fallas del equipo que inutilizan esas funciones.
- Establecer prioridades en los modos de falla.
- Definir las tareas de mantenimiento preventivo para modos de falla de alta prioridad.

El proyecto RCM duró de septiembre de 1988 a junio de 1994. Se consideró un total de 3 778 componentes en los 28 subsistemas. Al final del programa, en las políticas de mantenimiento preventivo se incluyeron más de 5 400 tareas para esos componentes. El costo de implementar este mantenimiento fue considerable: unos 30 000 dólares por sistema. Sin embargo, los beneficios superaron esos costos. Durante el periodo de 1990 a 1994, los registros indican una disminución apreciable en las fallas del equipo en esa planta. Además, un programa de mantenimiento basado en la confiabilidad puede traer otros beneficios, como los siguientes:

- Incrementar la disponibilidad de la planta.
- Optimizar los inventarios de refacciones.
- Identificar los modos de falla de componentes.
- Descubrir nuevos escenarios de falla de la planta.
- Adiestrar al personal técnico.
- Identificar componentes que aprovechan las estrategias del mantenimiento preventivo modificado.
- Identificar mejoras potenciales en el diseño.
- Mejorar la documentación.

Fox *et al*. (1994) informan algunas lecciones que dejó esta experiencia. Una es que para la organización interna de mantenimiento, es mejor dirigir el proceso, y no que lo haga una agencia externa. Con esto se evita el síndrome de “nosotros contra ellos”. En este caso, es probable también que la implementación tenga éxito. Se desplegó una lista de comprobación de análisis de costos para eliminar los modos de falla. Por último, el equipo desarrolló un sistema informático de base de datos relacional para facilitar las evaluaciones del RCM. Con este sistema se redujo en 50% el tiempo necesario para efectuar los análisis necesarios.

La lección que dejó este caso es que un programa de mantenimiento basado en la confiabilidad, diseñado e implementado cuidadosamente, puede producir grandes rendimientos en sistemas muy riesgosos.

constante de proporcionalidad igual a ϕ. Esto es, el tiempo para detectar el primer error tiene distribución exponencial con parámetro $N\phi$, el tiempo entre la detección del primer y segundo errores tiene distribución exponencial con parámetro $(N - 1)\phi$; y así sucesivamente.

Por consiguiente, a medida que se eliminan errores del programa, aumenta la cantidad de tiempo necesaria para detectar el siguiente error. Después de haber eliminado n errores, se habrán observado los valores de $T_1, T_2, \ldots, T_n$ que representan el tiempo entre las detecciones sucesivas. Con esas observaciones se estiman ϕ y N mediante el principio de máxima verosimilitud. Con base en estos estimados, se podría predecir con exactitud cuántas pruebas se necesitan para alcanzar cierto nivel de confiabilidad en el programa.

Shooman (1972) sugiere emplear una tasa normalizada de errores para medir el contenido de errores en el programa. Define a

$$p(t) = \text{Errores por cantidad total de instrucciones por mes de tiempo dedicado a corregir errores}$$

y deduce un modelo de confiabilidad basado en principios básicos. Demuestra cómo usar este modelo para establecer una relación funcional entre la cantidad de tiempo dedicada a corregir errores y la confiabilidad del programa.

Los trabajos de Jelinski y Moranda, y de Shooman, representan las bases de la teoría sobre la confiabilidad de software. Se han considerado ampliaciones de sus métodos. Sin embargo, queda por verse si esos métodos permiten obtener descripciones fieles del problema y si en último término ayudan a predecir el tiempo requerido para lograr un nivel aceptable de confiabilidad.

12.10 NOTAS HISTÓRICAS

Gran parte de la teoría de confiabilidad, las pruebas de tiempo de vida y las estrategias de mantenimiento tienen su raíz en la teoría actuarial desarrollada en las aseguradoras. Desde comienzos de siglo se cuenta con complicados modelos matemáticos que predicen las probabilidades de supervivencia. Lotka (1939) describe algunas de las relaciones entre los modelos de remplazo de equipo y los estudios actuariales. El trabajo de Weibull (1939 y 1951) es la base para el tema de la duración de los materiales sometidos a la fatiga.

El interés en los problemas de confiabilidad se difundió mucho más durante la Segunda Guerra Mundial, cuando se trató de comprender las leyes de fallas que dominan en los sistemas militares complejos. Durante la década de 1950 comenzaron a recibir gran atención los problemas sobre pruebas para establecer la duración y para determinar la confiabilidad de los misiles. En 1952, el Departamento de Defensa de Estados Unidos estableció el Grupo Asesor sobre Fiabilidad de Equipo Electrónico, que publicó su primer informe sobre confiabilidad en junio de 1957.

Los orígenes de los modelos específicos de remplazo por edad que se presentaron en este capítulo no son muy claros. Sin embargo, los modelos complicados datan de principios de la década de 1920 (véase Taylor, 1923, y Hotelling, 1925). Los modelos estocásticos de remplazo planeado presentados en la sección 7 forman la base de gran parte de la investigación en la teoría del remplazo, pero tampoco son claros los orígenes de esos modelos.

La sección sobre garantías se basa en el trabajo de Blischke y Scheuer (1975). Se pueden encontrar extensiones y correcciones de ese trabajo en Mamer (1982). Los lectores interesados en profundizar sus conocimientos deben consultar los excelentes textos de Barlow y Proschan (1965 y 1975) sobre modelos de confiabilidad, y de Gertsbakh (1977) sobre estrategias de mantenimiento. Los aspectos de la aplicación de modelos de mantenimiento se describen en Turban (1967) y en Mann (1976).

12.11 Resumen

El objetivo de este capítulo fue repasar la terminología y la metodología en torno a la teoría y aplicación de modelos de confiabilidad y de mantenimiento. La *teoría de confiabilidad* es una disciplina que ha recibido gran atención por parte de los matemáticos. Sin embargo, para nuestros fines las matemáticas no son lo único que interesa. Esos modelos tienen extremada utilidad en un entorno operativo, cuando se consideran asuntos como características de falla del equipo en funcionamiento, las estrategias de mantenimiento económicamente válidas y el valor de las garantías de producto y el de los contratos de servicio.

La complejidad del análisis depende de los supuestos que se hagan sobre la variable aleatoria T, que representa la vida de un solo artículo o equipo trabajando. La *función de distribución* de T, $F(t)$, es la probabilidad de que falle el artículo en el tiempo t o antes de éste ($P\{T \leq t\}$), mientras que la *función de confiabilidad* de T, $R(t)$, es la probabilidad de que el artículo falle después del tiempo t ($P\{T > t\}$). Una cantidad importante relacionada con esas funciones es la *función de tasa de falla,* $r(t)$, que es la relación $f(t)/R(t)$ de la función de densidad de probabilidades y la función de confiabilidad. Si Δt es suficientemente pequeño, el término $r(t)\Delta t$ se puede interpretar como la probabilidad condicional de que el artículo falle en las siguientes Δt unidades de tiempo, pues ha sobrevivido hasta el tiempo t.

La función de tasa de fallas permite tener mucha información sobre las características de envejecimiento del equipo en operación. En un ambiente manufacturero, cabría esperar que la mayor parte del equipo en funcionamiento tuviera una función creciente de tasa de fallas. Esto significa que es más probable que el artículo falle a medida que envejece. Se puede tener una función decreciente de tasa de fallas cuando la probabilidad de fallas tempranas se debe a piezas defectuosas de la población. La ley de probabilidad de *Weibull* se puede aplicar para describir las características de falla del equipo, con función creciente o decreciente de tasa de fallas.

Resulta interesante el caso en el que la función de tasa de fallas es constante. Tal caso da lugar a la *distribución exponencial* de la vida de un solo componente. La distribución exponencial es la única distribución continua que posee la *propiedad de amnesia.* Esto señala que la propiedad condicional de que un artículo que ha estado trabajando hasta el tiempo t falla en las siguientes unidades de tiempo, independientemente de t.

El *proceso de Poisson* describe el caso en el que un solo equipo falla siguiendo la distribución exponencial, y se le remplaza de inmediato al fallar. Cuando esto sucede, la cantidad de fallas en determinado tiempo tiene la distribución de Poisson, el tiempo entre fallas sucesivas tiene la distribución exponencial, y el tiempo para que se presenten n fallas tiene la distribución de Erlang.

En el capítulo se consideraron las funciones de confiabilidad de sistemas complejos de componentes. Se demostró cómo obtener las funciones de confiabilidad para *componentes en serie y en paralelo* a partir de las funciones de confiabilidad de los componentes individuales. También se describieron los sistemas K de N, que sólo funcionan si cuando menos funcionan K componentes.

Los aspectos de confiabilidad forman la base de los *modelos de mantenimiento* descritos en la segunda mitad del capítulo. Una medida importante del funcionamiento del sistema es la *disponibilidad,* que es la proporción del tiempo en que trabaja el equipo. Describimos los modelos deterministas de remplazo por edad, donde no se incluye en forma explícita la probabilidad de falla del equipo, y también los modelos estocásticos de remplazo por edad, que sí lo hacen. Los modelos estocásticos permiten remplazar el equipo antes de que falle. Esto es importante cuando los artículos tienen función creciente de la tasa de fallas, y fallas no planeadas que son más costosas que las fallas planeadas.

Por último, el capítulo terminó con una descripción del valor económico de las *garantías.* Una garantía es una promesa del vendedor al comprador de remplazar el artículo con uno nuevo si falla durante el periodo de garantía (*garantía de remplazo sin costo*), o bien de otorgar un descuento proporcional al tiempo restante de la garantía (o al desgaste) al comprar un nuevo artículo en el periodo de garantía (*garantía de prorrata*). Los asuntos relacionados con las

garantías y los contratos de servicio son parecidos, pero los modelos de contrato de servicio son mucho más complicados debido a la necesidad de incluir varios niveles de reparación.

Problemas adicionales sobre confiabilidad y facilidad de mantenimiento

42. Un gran productor nacional de electrodomésticos ha seguido de cerca la experiencia de los clientes con un tostador. En una encuesta de 5 000 clientes que compraron el tostador a principios de 1975 se reveló lo siguiente:

Año	Número de descomposturas
1975	188
1976	58
1977	63
1978	72
1979	54
1980	71

a) Usando estos datos, estime p_k = probabilidad de que un tostador falle en su k-ésimo año de funcionamiento, para $k = 1, \ldots, 6$.

b) ¿Cuál es la probabilidad de que un tostador dure cuando menos 6 años sin fallar, de acuerdo con estos datos?

c) La función discreta de tasa de fallas tiene la forma $r_k = p_k/R_{k-1}$, donde R_k es la probabilidad de que una unidad sobreviva el periodo k. Determine la función de tasa de fallas para los primeros 5 años de funcionamiento, a partir de los datos anteriores.

d) Suponga que usted compró un tostador a principios de 1984 y que todavía funciona al final de 1987. Si no ha cambiado en forma apreciable la confiabilidad, entre 1975 y 1984, use los resultados de la parte *e*) para obtener la probabilidad de que falle durante los dos primeros meses del año calendario 1988.

43. En una gran marquesina de un hotel y casino hay 6 000 focos. Cada uno falla totalmente al azar, y su vida promedio es 3 280 horas. Suponiendo que la marquesina permanece encendida en forma continua, y que los focos que se queman se remplazan de inmediato, ¿cuántos remplazos hay que hacer, en promedio, cada año?

44. El propietario del hotel mencionado en el problema 43 ha decidido que, para disminuir la cantidad de focos fundidos, remplazará los 6 000 focos al principio de cada año, además de remplazar los que se queman justo después de fallar. Comente la eficacia de esta estrategia.

45. El dueño del hotel mencionado en el problema 43 tiene tiempos difíciles y no remplaza los focos quemados. Observa que se quemó más de la mitad de los focos antes de la vida supuesta de 3 280 horas, y decide demandar al fabricante de focos por propaganda falsa. ¿Cree usted que ganará? (Sugerencia: ¿qué fracción de los focos se esperaría que fallaran antes de la vida promedio?)

46. Siguiendo con el ejemplo del problema 43, determine lo siguiente:

a) La proporción de focos que duran más de 2 años.

b) La probabilidad de que un foco tomado al azar falle en los tres primeros meses de funcionamiento.

c) La probabilidad de que un foco que ha durado 10 años falle en los siguientes 3 meses de operación.

47. Suponga que no se remplazan los focos del problema 43 en el momento en que se funden.

    a) ¿Qué fracción de los 6 000 focos se esperaría que fallaran en el primer año?

    b) ¿Qué fracción de los focos que sobrevivieron el primer año se esperaría que fallaran en el segundo?

    c) ¿Qué fracción de los focos que sobrevivieron al n-ésimo año se esperaría que fallaran en el año $n + 1$, para cualquier valor de $n = 1, 2, \ldots$?

    d) Con los resultados de la parte *c*), de los 6 000 bulbos originales, ¿cuántos se esperaría que fallaran en el cuarto año de funcionamiento?

48. El valor de la media de una variable aleatoria de Weibull se expresa con la fórmula

$$\mu = a^{-1/\beta}\,\Gamma(1 + 1/\beta),$$

donde Γ representa la función gamma. Esta función tiene la propiedad de que $\Gamma(k) = (k - 1)\Gamma(k - 1)$ para cualquier valor de $k > 1$, y además $\Gamma(1) = 1$. Observe que si k es entero, esto da como resultado $\Gamma(k) = (k - 1)!$ Si k no es entero, se debe recurrir a la definición recursiva de $\Gamma(k)$ junto con la tabla siguiente. Para valores de $1 \leq k \leq 2$, $\Gamma(k)$ es

k	**Γ(*k*)**	***k***	**Γ(*k*)**
1.00	1.0000	1.55	.8889
1.05	.9735	1.60	.8935
1.10	.9514	1.65	.9001
1.15	.9330	1.70	.9086
1.20	.9182	1.75	.9191
1.25	.9064	1.80	.9314
1.30	.8975	1.85	.9456
1.35	.8912	1.90	.9612
1.40	.8873	1.95	.9799
1.45	.8857	2.00	1.0000
1.50	.8862		

Por ejemplo, esta tabla se usaría como sigue: $\Gamma(3.6) = (2.6)\Gamma(2.6) = (2.6)(1.6)\Gamma(1.6) = (2.6)(1.6)(.8935) = 3.717$.

a) Calcule el tiempo esperado de falla para el ejemplo 12.4 acerca del equipo de copiado.

b) Calcule el tiempo esperado de falla para un equipo en funcionamiento, cuya ley de falla se da en el ejemplo 12.2.

c) Determine el tiempo medio a la falla para $\alpha = 1.35$ y $\beta = .20$.

d) Determine el tiempo medio a la falla para $\alpha = .90$ y $\beta = .45$.

49. Suponga que la propaganda indica que determinado foco tiene una vida promedio de 2 000 horas, y que se sabe que se apega a la ley exponencial de fallas. Para simplificar, suponga que se usa continuamente. Calcule la probabilidad de que el foco dure

    a) Más de 3 000 horas.

    b) Menos que 1 500 horas.

    c) Entre 2 000 y 2 500 horas.

50. Applicational Materials vende varios equipos que se usan en la fabricación de microprocesadores basados en silicio. En 1983, la empresa surtió 130 pedidos del modelo a5512. Suponga que las máquinas fallan de acuerdo con una ley de Weibull. En especial, la función $F(t)$ de distribución acumulada del tiempo hasta la falla de cualquiera de las máquinas se representa por

$$F(t) = 1 - e^{-0.0475t^{1.2}} \qquad \text{para toda } t \geq 0,$$

donde t está en años.

a) ¿Cuál es la función de tasa de fallas para este equipo?

b) De los 130 equipos originales vendidos en 1983, ¿cuántas máquinas se esperaría que no se descompusieran antes de enero de 1987? Para simplificar, suponga que todas se vendieron el primero de enero de 1983.

c) Con los resultados del inciso (*a*), estime la fracción de las máquinas que sobreviven 10 años de uso, y que se descomponen en el año 11 [o bien lo puede calcular directamente, si no tiene la respuesta al inciso (*a*)].

51. Una empresa tiene una flotilla de 10 taxis. Cada vez que uno se descompone, es reparado el mismo día. Suponga que las descomposturas de los coches suceden completamente al azar, con frecuencia de dos por año.

a) ¿Cuál es la probabilidad de que determinado taxi trabaje durante todo un año sin descomponerse?

b) ¿Cuál es la probabilidad de que toda la flota funcione un mes sin tener una descompostura?

c) En promedio, ¿cuántas descomposturas se esperarían en la flotilla en un periodo normal de 3 meses?

d) ¿Cuál es la probabilidad de que haya más de cinco descomposturas entre el 28 de noviembre y el Año Nuevo?

e) ¿Por qué razón podría ser demasiado baja la probabilidad obtenida en el inciso (*d*)?

52. Un juego de 30 luces de Navidad se dispone en un circuito en serie; esto significa que si uno de los focos se funde, toda la serie se apaga. Suponga que cada foco falla totalmente al azar, con una tasa de uno cada año. ¿Cuál es la probabilidad de que las luces funcionen desde el principio de la Nochebuena (24 de diciembre) hasta el final del Año Nuevo (1 de enero) sin fallar?

53. Cuesta $48 000 remplazar un equipo industrial, y esencialmente no tiene valor de recuperación. Durante los 5 primeros años de funcionamiento, los costos de mantenimiento fueron $8 000. Si la tasa de costos de mantenimiento es una función lineal del tiempo, ¿cuál es la edad óptima a la que se debe remplazar la máquina?

54. Para un automóvil de su propiedad, o uno del que le gustaría ser propietario, estime los valores correctos del costo de remplazo, la tasa de depreciación, el costo inicial de mantenimiento y la tasa a la que aumenta el costo de mantenimiento. Con base en estos datos, determine la cantidad óptima de años que debe usted esperar para cambiar de coche.

Apéndice 12-A

Glosario de notación sobre confiabilidad y facilidad de mantenimiento

a = Tasa de costo de mantenimiento por unidad de tiempo para el modelo simple de remplazo por edad. También se usa como parámetro de la función exponencial de costo de mantenimiento, para el modelo exponencial de remplazo por edad.

a_1 = Costo de remplazo de un artículo cuando falla. Se usa en el modelo de remplazo en bloque.

a_2 = Costo de remplazar todo el bloque de n_0 artículos. Se usa en el modelo de remplazo en bloque.

α = Parámetro de la función de probabilidad de Weibull.

b = Parámetro de la función exponencial de costo de mantenimiento, para el modelo exponencial de remplazo por edad.

β = Parámetro de la función de probabilidad de Weibull.

c = Parámetro de la función exponencial de valor de salvamento para el modelo exponencial de remplazo por edad.

c_1 = Costo de remplazar un artículo cuando falla. Se usa en el modelo de remplazo planeado bajo incertidumbre, para un solo artículo.

c_2 = Costo de remplazo de un artículo antes de que falle. Se usa en el modelo de remplazo planeado bajo incertidumbre, para un solo artículo.

C_1 = Costo de compra de un nuevo artículo con garantía de reposición sin costo.

C_2 = Costo de comprar un nuevo artículo con garantía de reposición de prorrata.

d = Un parámetro de la función exponencial de valor de recuperación, para el modelo exponencial de remplazo por edad.

$f(t)$ = Función de densidad de la variable aleatoria T.

$F(t) = P\{T \leq t\}$ = Función de distribución acumulada de T.

$F_K(t)$ = Función de distribución acumulada de la variable aleatoria T_k, el tiempo vida de un sistema K de N.

$F_P(t)$ = Función de distribución acumulada de la variable aleatoria T_p, el tiempo de vida de un sistema en paralelo.

$F_S(t)$ = Función de distribución acumulada de la variable aleatoria T_s, el tiempo de vida de un sistema en serie.

$G(t)$ = Costo promedio por unidad de tiempo para los modelos de remplazo por edad y de garantía.

K = Costo fijo de remplazo para modelos de remplazo por edad. También se usa para representar el costo de reposición de un artículo nuevo sin garantía, en los modelos de garantía.

λ = El parámetro de la ley exponencial de fallas; cantidad esperada de fallas por unidad de tiempo, cuando las fallas se presentan al azar.

TMEF = Tiempo medio entre fallas.

TPDR = Tiempo medio de reparación.

n_0 = Cantidad total de artículos en un bloque. Se usa en el modelo de remplazo en bloque.

$N(t)$ = Cantidad de fallas que suceden en el intervalo $(0, t)$, cuando las fallas siguen un proceso de Poisson. $N(t)$ tiene la distribución de Poisson con parámetro λt.

$r(t)$ = Función de tasa de fallas de T.

$R(t) = P\{T > t\}$ = Función de confiabilidad de T.

$R_K(t)$ = Función de confiabilidad de un sistema de componentes K de N.

$R_P(t)$ = Función de confiabilidad de un sistema de componentes en paralelo.

$R_S(t)$ = Función de confiabilidad de un sistema de componentes en serie.

T = Variable aleatoria que corresponde al tiempo de vida de una unidad en operación.

T_K = Tiempo de vida de un sistema de componentes K de N.

T_P = Tiempo de vida de un sistema de componentes en paralelo.

T_S = Tiempo de vida de un sistema de componentes en serie.

W_1 = Periodo de garantía, con garantía de remplazo sin costo.

W_2 = Periodo de garantía, con garantía de prorrata.

W_n = Tiempo para que sucedan n fallas, cuando las fallas siguen un proceso de Poisson. W_n tiene la distribución de Erlang, con parámetros λ y n.

Bibliografía

Amstadter, B. L. *Reliability Mathematics*. Nueva York: McGraw-Hill, 1971.

Barlow, R. E y F. Proschan. *Mathematical Theory of Reliability.* Nueva York: John Wiley & Sons, 1965.

Barlow, R. E. y F. Proschan. *Statistical Theory of Reliability and Life Testing.* Nueva York: Holt, Rinehart & Winston, 1975.

Blischke, W. R. y E. M. Scheuer, "Calculation of the Cost of Warranty Policies as a Function of Estimated Life Distributions". *Naval Research Logistics Quarterly* 22 (1975), pp. 681-96.

Davis, R. "As Complexity Rises, Tiny Flaws in Software Pose a Growing Threat". *The Wall Street Journal,* 28 de enero de 1987, p. 1.

DeGroot, M. H. *Probability and Statistics.* 2a. ed. Reading, MA: Addison- Wesley Publishing, 1986.

Fox, B. H., M. G. Snyder y A. M. Smith. "Reliability-Centered Maintenance Improves Operations at TMI Nuclear Plant". *Power Engineering,* noviembre de 1994, pp. 75-78.

Garvin, O. A. *Managing Quality.* Nueva York: The Free Press, 1988.

Gertsbakh, I. B. *Models of Preventive Maintenance.* Amsterdam: North Holland, 1977.

Hillier, F. S. y G. J. Lieberman. *Introduction to Operations Research.* Oakland, CA: Holden Day, 1986.

Hotelling, H. "A General Mathematical Theory of Depreciation". *Journal of the American Statistical Association* 20 (septiembre de 1925), pp. 340-53.

Jelinski, Z. y P. Moranda. "Software Reliability Research". En *Statistical Computer Performance Evaluation,* W. Freiberger, editor. Nueva York: Academic Press, 1972.

Lotka, A. J. "A Contribution to the Theory of Self-Renewing Aggregates with Special Reference to Industrial Replacement". *Annals of Mathematical Statistics* 10 (1939), pp. 1-25.

Lund, R. T. *Consumer Durables: Warranties, Service Contracts, and Alternatives.* Vol. 4, *Analyses of Consumer Product and Warranty Relationships.* Centro de Alternativas de Política, Instituto Tecnológico de Massachusetts, 1978.

Mamer, J. "Cost Analysis of Pro Rata and Free Replacement Warranties". *Naval Research Logistics Quarterly* 29 (1982), pp. 345-56.

Mann, L. L. *Maintenance Management. Lexington. MA:* Lexington Books, 1976.

McCall, J. J. "Maintenance Policies for Stochastically Failing Equipment: A Survey". *Management Science* 11 (1965), pp. 493-524.

O'Connor, P. D. T. *Practical Reliability Engineering.* 2a. ed., Nueva York: John Wiley & Sons, 1985.

Shooman, M. L. "Probabilistic Models for Software Reliability Prediction". *Statistical Computer Performance Evaluation.* W. Freiberger, editor, pp. 485-502. Nueva York: Academic Press, 1972.

Sivazlian, B. D. y L. E. Stanfel. *Analysis of Systems in Operations Research*. Englewood Cliffs, NJ: Prentice Hall, 1975.

Taylor, J. S. "A Statistical Theory of Depreciation". *Journal of the American Statistical Association,* diciembre de 1923, pp. 1010-23.

Terborgh, B. *Dynamic Equipment Policy.* Nueva York: McGraw-Hill, 1949.

Turban, E. "The Use of Mathematical Models in Plant Maintenance Decision Making". *Management Science* 13 (1967), pp. 20-27.

U.S. Nuclear Regulatory Commission. *Reactor Safety Study.* WASH-1400, NU REG 75-01, 1975.

Weibull, W. "A Statistical Theory of the Strength of Materials". *Ing. Vetenskaps Akad Handl.*, núm. 151 (1939).

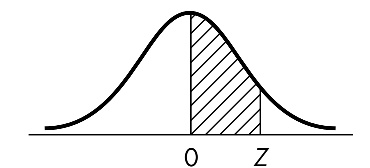
Weibull, W. "A Statistical Distribution Function of Wide Applicability". *Journal of Applied Mechanics* 18 (1951), pp. 293-97.

Wilkerson, J. J. "How to Manage Maintenance". *Harvard Business Review* 46 (1968), pp. 100-111.

Apéndice

Tablas

TABLA A-1 Áreas bajo la curva normal



Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.00	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.10	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.20	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.30	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.40	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.50	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.60	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.70	.2580	.2611	.2642	.2673	.2703	.2734	.2764	.2793	.2823	.2852
.80	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.90	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.00	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.10	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.20	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.30	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.40	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.50	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.60	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.70	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.80	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.90	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.00	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.10	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.20	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.30	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.40	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.50	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.60	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.70	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.80	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.90	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.00	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.10	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.20	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.30	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.40	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.50	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.60	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.70	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.80	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999

Los valores en el cuerpo de la tabla son las áreas entre la media y el valor de Z. Ejemplo: si se desea determinar el área bajo la curva normal estándar entre $Z = 0$ y $Z = 1.96$ se busca en el renglón $Z = 1.90$ y la columna .06 (porque $Z = 1.90 + .06 = 1.96$) y se lee .4750 en la intersección. Por comodidad, esta información también aparece en la tabla A-4.

Fuente: P. Billingsley, D.J. Croft, D.V. Huntsberger y C.J. Watson, *Statistical Inference for Managemente and Economics* (Boston: Allyn & Bacon, 1986). Reproducido con autorización.

TABLA A-2 Probabilidades binomiales acumuladas

$$P(X \leq r) = \sum_{k=0}^{r} \binom{n}{k} p^k (1-p)^{n-k}$$

donde X es la cantidad de éxitos en n intentos

$n = 5$

						p							
r	**.01**	**.05**	**.10**	**.20**	**.30**	**.40**	**.50**	**.60**	**.70**	**.80**	**.90**	**.95**	**.99**	***r***
0	0.9510	0.7738	0.5905	0.3277	0.1681	0.0778	0.0313	0.0102	0.0024	0.0003	0.0000	0.0000	0.0000	0
1	0.9990	0.9774	0.9185	0.7373	0.5282	0.3370	0.1875	0.0870	0.0308	0.0067	0.0005	0.0000	0.0000	1
2	1.0000	0.9988	0.9914	0.9421	0.8369	0.6826	0.5000	0.3174	0.1631	0.0579	0.0086	0.0012	0.0000	2
3	1.0000	1.0000	0.9995	0.9933	0.9692	0.9130	0.8125	0.6630	0.4718	0.2627	0.0815	0.0226	0.0010	3
4	1.0000	1.0000	1.0000	0.9997	0.9976	0.9898	0.9688	0.9222	0.8319	0.6723	0.4095	0.2262	0.0490	4

$n = 10$

						p							
r	**.01**	**.05**	**.10**	**.20**	**.30**	**.40**	**.50**	**.60**	**.70**	**.80**	**.90**	**.95**	**.99**	***r***
0	0.9044	0.5987	0.3487	0.1074	0.0282	0.0060	0.0010	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0
1	0.9957	0.9139	0.7361	0.3758	0.1493	0.0464	0.0107	0.0017	0.0001	0.0000	0.0000	0.0000	0.0000	1
2	0.9999	0.9885	0.9298	0.6778	0.3828	0.1673	0.0547	0.0123	0.0016	0.0001	0.0000	0.0000	0.0000	2
3	1.0000	0.9990	0.9872	0.8791	0.6496	0.3823	0.1719	0.0548	0.0106	0.0009	0.0000	0.0000	0.0000	3
4	1.0000	0.9999	0.9984	0.9672	0.8497	0.6331	0.3770	0.1662	0.0473	0.0064	0.0001	0.0000	0.0000	4
5	1.0000	1.0000	0.9999	0.9936	0.9527	0.8338	0.6230	0.3669	0.1503	0.0328	0.0016	0.0001	0.0000	5
6	1.0000	1.0000	1.0000	0.9991	0.9894	0.9452	0.8281	0.6177	0.3504	0.1209	0.0128	0.0010	0.0000	6
7	1.0000	1.0000	1.0000	0.9999	0.9984	0.9877	0.9453	0.8327	0.6172	0.3222	0.0702	0.0115	0.0001	7
8	1.0000	1.0000	1.0000	1.0000	0.9999	0.9983	0.9893	0.9536	0.8507	0.6242	0.2639	0.0861	0.0043	8
9	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9990	0.9940	0.9718	0.8926	0.6513	0.4013	0.0956	9

$n = 15$

						p							
r	**.01**	**.05**	**.10**	**.20**	**.30**	**.40**	**.50**	**.60**	**.70**	**.80**	**.90**	**.95**	**.99**	***r***
0	0.8601	0.4633	0.2059	0.0352	0.0047	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
1	0.9904	0.8290	0.5490	0.1671	0.0353	0.0052	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1
2	0.9996	0.9638	0.8159	0.3980	0.1268	0.0271	0.0037	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	2
3	1.0000	0.9945	0.9444	0.6482	0.2969	0.0905	0.0176	0.0019	0.0001	0.0000	0.0000	0.0000	0.0000	3
4	1.0000	0.9994	0.9873	0.8358	0.5155	0.2173	0.0592	0.0093	0.0007	0.0000	0.0000	0.0000	0.0000	4
5	1.0000	0.9999	0.9978	0.9389	0.7216	0.4032	0.1509	0.0338	0.0037	0.0001	0.0000	0.0000	0.0000	5
6	1.0000	1.0000	0.9997	0.9819	0.8689	0.6098	0.3036	0.0950	0.0152	0.0008	0.0000	0.0000	0.0000	6
7	1.0000	1.0000	1.0000	0.9958	0.9500	0.7869	0.5000	0.2131	0.0500	0.0042	0.0000	0.0000	0.0000	7
8	1.0000	1.0000	1.0000	0.9992	0.9848	0.9050	0.6964	0.3902	0.1311	0.0181	0.0003	0.0000	0.0000	8
9	1.0000	1.0000	1.0000	0.9999	0.9963	0.9662	0.8491	0.5968	0.2784	0.0611	0.0022	0.0001	0.0000	9
10	1.0000	1.0000	1.0000	1.0000	0.9993	0.9907	0.9408	0.7827	0.4845	0.1642	0.0127	0.0006	0.0000	10
11	1.0000	1.0000	1.0000	1.0000	0.9999	0.9981	0.9824	0.9095	0.7031	0.3518	0.0556	0.0055	0.0000	11
12	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9963	0.9729	0.8732	0.6020	0.1841	0.0362	0.0004	12
13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9948	0.9647	0.8329	0.4510	0.1710	0.0096	13
14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9953	0.9648	0.7941	0.5367	0.1399	14

TABLA A-2 ***(continuación)***

n = 20

						p							
r	**.01**	**.05**	**.10**	**.20**	**.30**	**.40**	**.50**	**.60**	**.70**	**.80**	**.90**	**.95**	**.99**	**r**
0	0.8179	0.3585	0.1216	0.0115	0.0008	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
1	0.9831	0.7358	0.3917	0.0692	0.0076	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1
2	0.9990	0.9245	0.6769	0.2061	0.0355	0.0036	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	2
3	1.0000	0.9841	0.8670	0.4114	0.1071	0.0160	0.0013	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	3
4	1.0000	0.9974	0.9568	0.6296	0.2375	0.0510	0.0059	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	4
5	1.0000	0.9997	0.9887	0.8042	0.4164	0.1256	0.0207	0.0016	0.0000	0.0000	0.0000	0.0000	0.0000	5
6	1.0000	1.0000	0.9976	0.9133	0.6080	0.2500	0.0577	0.0065	0.0003	0.0000	0.0000	0.0000	0.0000	6
7	1.0000	1.0000	0.9996	0.9679	0.7723	0.4159	0.1316	0.0210	0.0013	0.0000	0.0000	0.0000	0.0000	7
8	1.0000	1.0000	0.9999	0.9900	0.8867	0.5956	0.2517	0.0565	0.0051	0.0001	0.0000	0.0000	0.0000	8
9	1.0000	1.0000	1.0000	0.9974	0.9520	0.7553	0.4119	0.1275	0.0171	0.0006	0.0000	0.0000	0.0000	9
10	1.0000	1.0000	1.0000	0.9994	0.9829	0.8725	0.5881	0.2447	0.0480	0.0026	0.0000	0.0000	0.0000	10
11	1.0000	1.0000	1.0000	0.9999	0.9949	0.9435	0.7483	0.4044	0.1133	0.0100	0.0001	0.0000	0.0000	11
12	1.0000	1.0000	1.0000	1.0000	0.9987	0.9790	0.8684	0.5841	0.2277	0.0321	0.0004	0.0000	0.0000	12
13	1.0000	1.0000	1.0000	1.0000	0.9997	0.9935	0.9423	0.7500	0.3920	0.0867	0.0024	0.0000	0.0000	13
14	1.0000	1.0000	1.0000	1.0000	1.0000	0.9984	0.9793	0.8744	0.5836	0.1958	0.0113	0.0003	0.0000	14
15	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9941	0.9490	0.7625	0.3704	0.0432	0.0026	0.0000	15
16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9987	0.9840	0.8929	0.5886	0.1330	0.0159	0.0000	16
17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9964	0.9645	0.7939	0.3231	0.0755	0.0010	17
18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9924	0.9308	0.6083	0.2642	0.0169	18
19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9992	0.9885	0.8784	0.6415	0.1821	19

n = 25

						p							
r	**.01**	**.05**	**.10**	**.20**	**.30**	**.40**	**.50**	**.60**	**.70**	**.80**	**.90**	**.95**	**.99**	**r**
0	0.7778	0.2774	0.0718	0.0038	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
1	0.9742	0.6424	0.2712	0.0274	0.0016	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1
2	0.9980	0.8729	0.5371	0.0982	0.0090	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	2
3	0.9999	0.9659	0.7636	0.2340	0.0332	0.0024	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	3
4	1.0000	0.9928	0.9020	0.4207	0.0905	0.0095	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	4
5	1.0000	0.9988	0.9666	0.6167	0.1935	0.0294	0.0020	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	5
6	1.0000	0.9998	0.9905	0.7800	0.3407	0.0736	0.0073	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	6
7	1.0000	1.0000	0.9977	0.8909	0.5118	0.1536	0.0216	0.0012	0.0000	0.0000	0.0000	0.0000	0.0000	7
8	1.0000	1.0000	0.9995	0.9532	0.6769	0.2735	0.0539	0.0043	0.0001	0.0000	0.0000	0.0000	0.0000	8
9	1.0000	1.0000	0.9999	0.9827	0.8106	0.4246	0.1148	0.0132	0.0005	0.0000	0.0000	0.0000	0.0000	9
10	1.0000	1.0000	1.0000	0.9944	0.9022	0.5858	0.2122	0.0344	0.0018	0.0000	0.0000	0.0000	0.0000	10
11	1.0000	1.0000	1.0000	0.9985	0.9558	0.7323	0.3450	0.0778	0.0060	0.0001	0.0000	0.0000	0.0000	11
12	1.0000	1.0000	1.0000	0.9996	0.9825	0.8462	0.5000	0.1538	0.0175	0.0004	0.0000	0.0000	0.0000	12
13	1.0000	1.0000	1.0000	0.9999	0.9940	0.9222	0.6550	0.2677	0.0442	0.0015	0.0000	0.0000	0.0000	13
14	1.0000	1.0000	1.0000	1.0000	0.9982	0.9656	0.7878	0.4142	0.0978	0.0056	0.0000	0.0000	0.0000	14
15	1.0000	1.0000	1.0000	1.0000	0.9995	0.9868	0.8852	0.5754	0.1894	0.0173	0.0001	0.0000	0.0000	15
16	1.0000	1.0000	1.0000	1.0000	0.9999	0.9957	0.9461	0.7265	0.3231	0.0468	0.0005	0.0000	0.0000	16
17	1.0000	1.0000	1.0000	1.0000	1.0000	0.9988	0.9784	0.8464	0.4882	0.1091	0.0023	0.0000	0.0000	17
18	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9927	0.9264	0.6593	0.2200	0.0095	0.0002	0.0000	18
19	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9980	0.9706	0.8065	0.3833	0.0334	0.0012	0.0000	19
20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9905	0.9095	0.5793	0.0980	0.0072	0.0000	20
21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9976	0.9668	0.7660	0.2364	0.0341	0.0001	21
22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9996	0.9910	0.9018	0.4629	0.1271	0.0020	22
23	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9984	0.9726	0.7288	0.3576	0.0258	23
24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9962	0.9282	0.7226	0.2222	24

Fuente: P. Billingsley, D.J. Croft, D.V. Huntsberger, y C.J. Watson, *Statistical Inference for Management and Economics* (Boston: Allyn & Bacon, 1986).
Reproducido con autorización.

TABLA A-3 Probabilidades de Poisson acumuladas

$$\sum_{x=x'}^{\infty} \frac{e^{-m}m^x}{x!}$$

x'	m 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.0952	.1813	.2592	.3297	.3935	.4512	.5034	.5507	.5934	.6321
2	.0047	.0175	.0369	.0616	.0902	.1219	.1558	.1912	.2275	.2642
3	.0002	.0011	.0036	.0079	.0144	.0231	.0341	.0474	.0629	.0803
4	.0000	.0001	.0003	.0008	.0018	.0034	.0058	.0091	.0135	.0190
5	.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0014	.0023	.0037
6	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0006
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

x'	m 1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.6671	.6988	.7275	.7534	.7769	.7981	.8173	.8347	.8504	.8647
2	.3010	.3374	.3732	.4082	.4422	.4751	.5068	.5372	.5663	.5940
3	.0996	.1205	.1429	.1665	.1912	.2166	.2428	.2694	.2963	.3233
4	.0257	.0338	.0431	.0537	.0656	.0788	.0932	.1087	.1253	.1429
5	.0054	.0077	.0107	.0143	.0186	.0237	.0296	.0364	.0441	.0527
6	.0010	.0015	.0022	.0032	.0045	.0060	.0080	.0104	.0132	.0166
7	.0001	.0003	.0004	.0006	.0009	.0013	.0019	.0026	.0034	.0045
8	.0000	.0000	.0001	.0001	.0002	.0003	.0004	.0006	.0008	.0011
9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0002

x'	m 2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.8775	.8892	.8997	.9093	.9179	.9257	.9328	.9392	.9450	.9502
2	.6204	.6454	.6691	.6916	.7127	.7326	.7513	.7689	.7854	.8009
3	.3504	.3773	.4040	.4303	.4562	.4816	.5064	.5305	.5540	.5768
4	.1614	.1806	.2007	.2213	.2424	.2640	.2859	.3081	.3304	.3528
5	.0621	.0725	.0838	.0959	.1088	.1226	.1371	.1523	.1682	.1847
6	.0204	.0249	.0300	.0357	.0420	.0490	.0567	.0651	.0742	.0839
7	.0059	.0075	.0094	.0116	.0142	.0172	.0206	.0244	.0287	.0335
8	.0015	.0020	.0026	.0033	.0042	.0053	.0066	.0081	.0099	.0119
9	.0003	.0005	.0006	.0009	.0011	.0015	.0019	.0024	.0031	.0038
10	.0001	.0001	.0001	.0002	.0003	.0004	.0005	.0007	.0009	.0011
11	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

TABLA A-3 (*continuación*)

	m									
x′	**3.1**	**3.2**	**3.3**	**3.4**	**3.5**	**3.6**	**3.7**	**3.8**	**3.9**	**4.0**
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9550	.9592	.9631	.9666	.9698	.9727	.9753	.9776	.9798	.9817
2	.8153	.8288	.8414	.8532	.8641	.8743	.8838	.8926	.9008	.9084
3	.5988	.6201	.6406	.6603	.6792	.6973	.7146	.7311	.7469	.7619
4	.3752	.3975	.4197	.4416	.4634	.4848	.5058	.5265	.5468	.5665
5	.2018	.2194	.2374	.2558	.2746	.2936	.3128	.3322	.3516	.3712
6	.0943	.1054	.1171	.1295	.1424	.1559	.1699	.1844	.1994	.2149
7	.0388	.0446	.0510	.0579	.0653	.0733	.0818	.0909	.1005	.1107
8	.0142	.0168	.0198	.0231	.0267	.0308	.0352	.0401	.0454	.0511
9	.0047	.0057	.0069	.0083	.0099	.0117	.0137	.0160	.0185	.0214
10	.0014	.0018	.0022	.0027	.0033	.0040	.0048	.0058	.0069	.0081
11	.0004	.0005	.0006	.0008	.0010	.0013	.0016	.0019	.0023	.0028
12	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0007	.0009
13	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002	.0003
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

	m									
x′	**4.1**	**4.2**	**4.3**	**4.4**	**4.5**	**4.6**	**4.7**	**4.8**	**4.9**	**5.0**
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9834	.9850	.9864	.9877	.9889	.9899	.9909	.9918	.9926	.9933
2	.9155	.9220	.9281	.9337	.9389	.9437	.9482	.9523	.9561	.9596
3	.7762	.7898	.8026	.8149	.8264	.8374	.8477	.8575	.8667	.8753
4	.5858	.6046	.6228	.6406	.6577	.6743	.6903	.7058	.7207	.7350
5	.3907	.4102	.4296	.4488	.4679	.4868	.5054	.5237	.5418	.5595
6	.2307	.2469	.2633	.2801	.2971	.3142	.3316	.3490	.3665	.3840
7	.1214	.1325	.1442	.1564	.1689	.1820	.1954	.2092	.2233	.2378
8	.0573	.0639	.0710	.0786	.0866	.0951	.1040	.1133	.1231	.1334
9	.0245	.0279	.0317	.0358	.0403	.0451	.0503	.0558	.0618	.0681
10	.0095	.0111	.0129	.0149	.0171	.0195	.0222	.0251	.0283	.0318
11	.0034	.0041	.0048	.0057	.0067	.0078	.0090	.0104	.0120	.0137
12	.0011	.0014	.0017	.0020	.0024	.0029	.0034	.0040	.0047	.0055
13	.0003	.0004	.0005	.0007	.0008	.0010	.0012	.0014	.0017	.0020
14	.0001	.0001	.0002	.0002	.0003	.0003	.0004	.0005	.0006	.0007
15	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0002	.0002
16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

	m									
x′	**5.1**	**5.2**	**5.3**	**5.4**	**5.5**	**5.6**	**5.7**	**5.8**	**5.9**	**6.0**
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9939	.9945	.9950	.9955	.9959	.9963	.9967	.9970	.9973	.9975
2	.9628	.9658	.9686	.9711	.9734	.9756	.9776	.9794	.9811	.9826
3	.8835	.8912	.8984	.9052	.9116	.9176	.9232	.9285	.9334	.9380
4	.7487	.7619	.7746	.7867	.7983	.8094	.8200	.8300	.8396	.8488
5	.5769	.5939	.6105	.6267	.6425	.6579	.6728	.6873	.7013	.7149
6	.4016	.4191	.4365	.4539	.4711	.4881	.5050	.5217	.5381	.5543
7	.2526	.2676	.2829	.2983	.3140	.3297	.3456	.3616	.3776	.3937
8	.1440	.1551	.1665	.1783	.1905	.2030	.2159	.2290	.2424	.2560
9	.0748	.0819	.0894	.0974	.1056	.1143	.1234	.1328	.1426	.1528

(*continúa*)

TABLA A-3 (*continuación*)

	m									
x′	**5.1**	**5.2**	**5.3**	**5.4**	**5.5**	**5.6**	**5.7**	**5.8**	**5.9**	**6.0**
10	.0356	.0397	.0441	.0488	.0538	.0591	.0648	.0708	.0772	.0839
11	.0156	.0177	.0200	.0225	.0253	.0282	.0314	.0349	.0386	.0426
12	.0063	.0073	.0084	.0096	.0110	.0125	.0141	.0160	.0179	.0201
13	.0024	.0028	.0033	.0038	.0045	.0051	.0059	.0068	.0078	.0088
14	.0008	.0010	.0012	.0014	.0017	.0020	.0023	.0027	.0031	.0036
15	.0003	.0003	.0004	.0005	.0006	.0007	.0009	.0010	.0012	.0014
16	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0004	.0004	.0005
17	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0002
18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

	m									
x′	**6.1**	**6.2**	**6.3**	**6.4**	**6.5**	**6.6**	**6.7**	**6.8**	**6.9**	**7.0**
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9978	.9980	.9982	.9983	.9985	.9986	.9988	.9989	.9990	.9991
2	.9841	.9854	.9866	.9877	.9887	.9897	.9905	.9913	.9920	.9927
3	.9423	.9464	.9502	.9537	.9570	.9600	.9629	.9656	.9680	.9704
4	.8575	.8658	.8736	.8811	.8882	.8948	.9012	.9072	.9129	.9182
5	.7281	.7408	.7531	.7649	.7763	.7873	.7978	.8080	.8177	.8270
6	.5702	.5859	.6012	.6163	.6310	.6453	.6594	.6730	.6863	.6993
7	.4098	.4258	.4418	.4577	.4735	.4892	.5047	.5201	.5353	.5503
8	.2699	.2840	.2983	.3127	.3272	.3419	.3567	.3715	.3864	.4013
9	.1633	.1741	.1852	.1967	.2084	.2204	.2327	.2452	.2580	.2709
10	.0910	.0984	.1061	.1142	.1226	.1314	.1404	.1498	.1505	.1695
11	.0469	.0514	.0563	.0614	.0668	.0726	.0786	.0849	.0916	.0985
12	.0224	.0250	.0277	.0307	.0339	.0373	.0409	.0448	.0490	.0534
13	.0100	.0113	.0127	.0143	.0160	.0179	.0199	.0221	.0245	.0270
14	.0042	.0048	.0055	.0063	.0071	.0080	.0091	.0102	.0115	.0128
15	.0016	.0019	.0022	.0026	.0030	.0034	.0039	.0044	.0050	.0057
16	.0006	.0007	.0008	.0010	.0012	.0014	.0016	.0018	.0021	.0024
17	.0002	.0003	.0003	.0004	.0004	.0005	.0006	.0007	.0008	.0010
18	.0001	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0003	.0004
19	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001

	m									
x′	**7.1**	**7.2**	**7.3**	**7.4**	**7.5**	**7.6**	**7.7**	**7.8**	**7.9**	**8.0**
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9992	.9993	.9993	.9994	.9994	.9995	.9995	.9996	.9996	.9997
2	.9933	.9939	.9944	.9949	.9953	.9957	.9961	.9964	.9967	.9970
3	.9725	.9745	.9764	.0781	.9797	.9812	.9826	.9839	.9851	.9862
4	.9233	.9281	.9326	.9368	.9409	.9446	.9482	.9515	.9547	.9576
5	.8359	.8445	.8527	.8605	.8679	.8751	.8819	.8883	.8945	.9004
6	.7119	.7241	.7360	.7474	.7586	.7693	.7797	.7897	.7994	.8088
7	.5651	.5796	.5940	.6080	.6218	.6354	.6486	.6616	.6743	.6866
8	.4162	.4311	.4459	.4607	.4754	.4900	.5044	.5188	.5330	.5470
9	.2840	.2973	.3108	.3243	.3380	.3518	.3657	.3796	.3935	.4075

TABLA A-3 (*continuación*)

	m									
x′	**7.1**	**7.2**	**7.3**	**7.4**	**7.5**	**7.6**	**7.7**	**7.8**	**7.9**	**8.0**
10	.1798	.1904	.2012	.2123	.2236	.2351	.2469	.2589	.2710	.2834
11	.1058	.1133	.1212	.1293	.1378	.1465	.1555	.1648	.1743	.1841
12	.0580	.0629	.0681	.0735	.0792	.0852	.0915	.0980	.1048	.1119
13	.0297	.0327	.0358	.0391	.0427	.0464	.0504	.0546	.0591	.0638
14	.0143	.0159	.0176	.0195	.0216	.0238	.0261	.0286	.0313	.0342
15	.0065	.0073	.0082	.0092	.0103	.0114	.0127	.0141	.0156	.0173
16	.0028	.0031	.0036	.0041	.0046	.0052	.0059	.0066	.0074	.0082
17	.0011	.0013	.0015	.0017	.0020	.0022	.0026	.0029	.0033	.0037
18	.0004	.0005	.0006	.0007	.0008	.0009	.0011	.0012	.0014	.0016
19	.0002	.0002	.0002	.0003	.0003	.0004	.0004	.0005	.0006	.0006
20	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002	.0002	.0003
21	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001

	m									
x′	**8.1**	**8.2**	**8.3**	**8.4**	**8.5**	**8.6**	**8.7**	**8.8**	**8.9**	**9.0**
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9997	.9997	.9998	.9998	.9998	.9998	.9998	.9998	.9999	.9999
2	.9972	.9975	.9977	.9979	.9981	.9982	.9984	.9985	.9987	.9988
3	.9873	.9882	.9891	.9900	.9907	.9914	.9921	.9927	.9932	.9938
4	.9604	.9630	.9654	.9677	.9699	.9719	.9738	.9756	.9772	.9788
5	.9060	.9113	.9163	.9211	.9256	.9299	.9340	.9379	.9416	.9450
6	.8178	.8264	.8347	.8427	.8504	.8578	.8648	.8716	.8781	.8843
7	.6987	.7104	.7219	.7330	.7438	.7543	.7645	.7744	.7840	.7932
8	.5609	.5746	.5881	.6013	.6144	.6272	.6398	.6522	.6643	.6761
9	.4214	.4353	.4493	.4631	.4769	.4906	.5042	.5177	.5311	.5443
10	.2959	.3085	.3212	.3341	.3470	.3600	.3731	.3863	.3994	.4126
11	.1942	.2045	.2150	.2257	.2366	.2478	.2591	.2706	.2822	.2940
12	.1193	.1269	.1348	.1429	.1513	.1600	.1689	.1780	.1874	.1970
13	.0687	.0739	.0793	.0850	.0909	.0971	.1035	.1102	.1171	.1242
14	.0372	.0405	.0439	.0476	.0514	.0555	.0597	.0642	.0689	.0739
15	.0190	.0209	.0229	.0251	.0274	.0299	.0325	.0353	.0383	.0415
16	.0092	.0102	.0113	.0125	.0138	.0152	.0168	.0184	.0202	.0220
17	.0042	.0047	.0053	.0059	.0066	.0074	.0082	.0091	.0101	.0111
18	.0018	.0021	.0023	.0027	.0030	.0034	.0038	.0043	.0048	.0053
19	.0008	.0009	.0010	.0011	.0013	.0015	.0017	.0019	.0022	.0024
20	.0003	.0003	.0004	.0005	.0005	.0006	.0007	.0008	.0009	.0011
21	.0001	.0001	.0002	.0002	.0002	.0002	.0003	.0003	.0004	.0004
22	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002
23	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

	m									
x′	**9.1**	**9.2**	**9.3**	**9.4**	**9.5**	**9.6**	**9.7**	**9.8**	**9.9**	**10**
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	1.0000	1.0000
2	.9989	.9990	.9991	.9991	.9992	.9993	.9993	.9994	.9995	.9995
3	.9942	.9947	.9951	.9955	.9958	.9962	.9965	.9967	.9970	.9972
4	.9802	.9816	.9828	.9840	.9851	.9862	.9871	.9880	.9889	.9897

(*continúa*)

TABLA A-3 ***(continuación)***

	m									
x′	**9.1**	**9.2**	**9.3**	**9.4**	**9.5**	**9.6**	**9.7**	**9.8**	**9.9**	**10**
5	.9483	.9514	.9544	.9571	.9597	.9622	.9645	.9667	.9688	.9707
6	.8902	.8959	.9014	.9065	.9115	.9162	.9207	.9250	.9290	.9329
7	.8022	.8108	.8192	.8273	.8351	.8426	.8498	.8567	.8634	.8699
8	.6877	.6990	.7101	.7208	.7313	.7416	.7515	.7612	.7706	.7798
9	.5574	.5704	.5832	.5958	.6082	.6204	.6324	.6442	.6558	.6672
10	.4258	.4389	.4521	.4651	.4782	.4911	.5040	.5168	.5295	.5421
11	.3059	.3180	.3301	.3424	.3547	.3671	.3795	.3920	.4045	.4170
12	.2068	.2168	.2270	.2374	.2480	.2588	.2697	.2807	.2919	.3032
13	.1316	.1393	.1471	.1552	.1636	.1721	.1809	.1899	.1991	.2084
14	.0790	.0844	.0900	.0958	.1019	.1081	.1147	.1214	.1284	.1355
15	.0448	.0483	.0520	.0559	.0600	.0643	.0688	.0735	.0784	.0835
16	.0240	.0262	.0285	.0309	.0335	.0362	.0391	.0421	.0454	.0487
17	.0122	.0135	.0148	.0162	.0177	.0194	.0211	.0230	.0249	.0270
18	.0059	.0066	.0073	.0081	.0089	.0098	.0108	.0119	.0130	.0143
19	.0027	.0031	.0034	.0038	.0043	.0048	.0053	.0059	.0065	.0072
20	.0012	.0014	.0015	.0017	.0020	.0022	.0025	.0028	.0031	.0035
21	.0005	.0006	.0007	.0008	.0009	.0010	.0011	.0013	.0014	.0016
22	.0002	.0002	.0003	.0003	.0004	.0004	.0005	.0005	.0006	.0007
23	.0001	.0001	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0003
24	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001

	m									
x′	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	.9998	.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	.9988	.9995	.9998	.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	.9951	.9977	.9990	.9995	.9998	.9999	1.0000	1.0000	1.0000	1.0000
5	.9849	.9924	.9963	.9982	.9991	.9996	.9998	.9999	1.0000	1.0000
6	.9625	.9797	.9893	.9945	.9972	.9986	.9993	.9997	.9998	.9999
7	.9214	.9542	.9741	.9858	.9924	.9960	.9979	.9990	.9995	.9997
8	.8568	.9105	.9460	.9684	.9820	.9900	.9946	.9971	.9985	.9992
9	.7680	.8450	.9002	.9379	.9626	.9780	.9874	.9929	.9961	.9979
10	.6595	.7576	.8342	.8906	.9301	.9567	.9739	.9846	.9911	.9950
11	.5401	.6528	.7483	.8243	.8815	.9226	.9509	.9696	.9817	.9892
12	.4207	.5384	.6468	.7400	.8152	.8730	.9153	.9451	.9653	.9786
13	.3113	.4240	.5369	.6415	.7324	.8069	.8650	.9083	.9394	.9610
14	.2187	.3185	.4270	.5356	.6368	.7255	.7991	.8574	.9016	.9339
15	.1460	.2280	.3249	.4296	.5343	.6325	.7192	.7919	.8503	.8951
16	.0926	.1556	.2364	.3306	.4319	.5333	.6285	.7133	.7852	.8435
17	.0559	.1013	.1645	.2441	.3359	.4340	.5323	.6250	.7080	.7789
18	.0322	.0630	.1095	.1728	.2511	.3407	.4360	.5314	.6216	.7030
19	.0177	.0374	.0698	.1174	.1805	.2577	.3450	.4378	.5305	.6186

TABLA A-3 (*continuación*)

	m									
x′	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
20	.0093	.0213	.0427	.0765	.1248	.1878	.2637	.3491	.4394	.5297
21	.0047	.0116	.0250	.0479	.0830	.1318	.1945	.2693	.3528	.4409
22	.0023	.0061	.0141	.0288	.0531	.0892	.1385	.2009	.2745	.3563
23	.0010	.0030	.0076	.0167	.0327	.0582	.0953	.1449	.2069	.2794
24	.0005	.0015	.0040	.0093	.0195	.0367	.0633	.1011	.1510	.2125
25	.0002	.0007	.0020	.0050	.0112	.0223	.0406	.0683	.1067	.1568
26	.0001	.0003	.0010	.0026	.0062	.0131	.0252	.0446	.0731	.1122
27	.0000	.0001	.0005	.0013	.0033	.0075	.0152	.0282	.0486	.0779
28	.0000	.0001	.0002	.0006	.0017	.0041	.0088	.0173	.0313	.0525
29	.0000	.0000	.0001	.0003	.0009	.0022	.0050	.0103	.0195	.0343
30	.0000	.0000	.0000	.0001	.0004	.0011	.0027	.0059	.0118	.0218
31	.0000	.0000	.0000	.0001	.0002	.0006	.0014	.0033	.0070	.0135
32	.0000	.0000	.0000	.0000	.0001	.0003	.0007	.0018	.0040	.0081
33	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0010	.0022	.0047
34	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0012	.0027
35	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0015
36	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008
37	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004
38	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
39	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
40	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

Fuente: *CRC Standard Mathematical Tables*.16a edición, The Chemical Rubber Company, 1968.

TABLA A-4 Distribución normal de probabilidades y esperanzas parciales

Variable estandarizada	Probabilidades		Esperanzas parciales	
z	$F(z)$	$1 - F(z)$	$L(z)$	$L(-z)$
.00	.5000	.5000	.3989	.3989
.01	.5040	.4960	.3940	.4040
.02	.5080	.4920	.3890	.4090
.03	.5120	.4880	.3841	.4141
.04	.5160	.4840	.3793	.4193
.05	.5200	.4800	.3744	.4244
.06	.5239	.4761	.3697	.4297
.07	.5279	.4721	.3649	.4349
.08	.5319	.4681	.3602	.4402
.09	.5359	.4641	.3556	.4456
.10	.5398	.4602	.3509	.4509
.11	.5438	.4562	.3464	.4564
.12	.5478	.4522	.3418	.4618
.13	.5517	.4483	.3373	.4673
.14	.5557	.4443	.3328	.4728
.15	.5596	.4404	.3284	.4784
.16	.5636	.4364	.3240	.4840
.17	.5685	.4325	.3197	.4897
.18	.5714	.4286	.3154	.4954
.19	.5753	.4247	.3111	.5011
.20	.5793	.4207	.3069	.5069
.21	.5832	.4168	.3027	.5127
.22	.5871	.4129	.3027	.5186
.23	.5910	.4090	.2944	.5244
.24	.5948	.4052	.2904	.5304
.25	.5987	.4013	.2863	.5363
.26	.6026	.3974	.2824	.5424
.27	.6064	.3936	.2784	.5484
.28	.6103	.3897	.2745	.5545
.29	.6141	.3859	.2706	.5606
.30	.6179	.3821	.2668	.5668
.31	.6217	.3783	.2630	.5730
.32	.6255	.3745	.2592	.5792
.33	.6293	.3707	.2555	.5855
.34	.6331	.3669	.2518	.5918
.35	.6368	.3632	.2481	.5981
.36	.6406	.3594	.2445	.6045
.37	.6443	.3557	.2409	.6109
.38	.6480	.3520	.2374	.6174
.39	.6517	.3483	.2339	.6239
.40	.6554	.3446	.2304	.6304
.41	.6591	.3409	.2270	.6370
.42	.6628	.3372	.2236	.6436
.43	.6664	.3336	.2203	.6503
.44	.6700	.3300	.2169	.6569
.45	.6736	.3264	.2137	.6637
.46	.6772	.3228	.2104	.6704
.47	.6808	.3192	.2072	.6772
.48	.6844	.3156	.2040	.6840
.49	.6879	.3121	.2009	.6909
.50	.6915	.3085	.1978	.6978

TABLA A-4 ***(continuación)***

Variable estandarizada	Probabilidades		Esperanzas parciales	
z	$F(z)$	$1 - F(z)$	$L(z)$	$L(-z)$
.51	.6950	.3050	.1947	.7047
.52	.6985	.3015	.1917	.7117
.53	.7019	.2981	.1887	.7187
.54	.7054	.2946	.1857	.7257
.55	.7088	.2912	.1828	.7328
.56	.7123	.2877	.1799	.7399
.57	.7157	.2843	.1771	.7471
.58	.7190	.2810	.1742	.7542
.59	.7224	.2776	.1714	.7614
.60	.7257	.2743	.1687	.7687
.61	.7291	.2709	.1659	.7759
.62	.7324	.2676	.1633	.7833
.63	.7357	.2643	.1606	.7906
.64	.7389	.2611	.1580	.7980
.65	.7422	.2578	.1554	.8054
.66	.7454	.2546	.1528	.8128
.67	.7486	.2514	.1503	.8203
.68	.7517	.2483	.1478	.8278
.69	.7549	.2451	.1453	.8353
.70	.7580	.2420	.1429	.8429
.71	.7611	.2389	.1405	.8505
.72	.7642	.2358	.1381	.8581
.73	.7673	.2327	.1358	.8658
.74	.7703	.2297	.1334	.8734
.75	.7733	.2267	.1312	.8812
.76	.7764	.2236	.1289	.8889
.77	.7793	.2207	.1267	.8967
.78	.7823	.2177	.1245	.9045
.79	.7852	.2148	.1223	.9123
.80	.7881	.2119	.1202	.9202
.81	.7910	.2090	.1181	.9281
.82	.7939	.2061	.1160	.9360
.83	.7967	.2033	.1140	.9440
.84	.7996	.2004	.1120	.9520
.85	.8023	.1977	.1100	.9600
.86	.8051	.1949	.1080	.9680
.87	.8067	.1922	.1061	.9761
.88	.8106	.1894	.1042	.9842
.89	.8133	.1867	.1023	.9923
.90	.8159	.1841	.1004	1.0004
.91	.8186	.1814	.0986	1.0086
.92	.8212	.1788	.0968	1.0168
.93	.8238	.1762	.0955	1.0250
.94	.8264	.1736	.0953	1.0330
.95	.8289	.1711	.0916	1.0416
.96	.8315	.1685	.0899	1.0499
.97	.8340	.1660	.0882	1.0582
.98	.8365	.1635	.0865	1.0665
.99	.8389	.1611	.0849	1.0749
1.00	.8413	.1587	.0833	1.0833
1.01	.8438	.1562	.0817	1.0917
1.02	.8461	.1539	.0802	1.1002
1.03	.8485	.1515	.0787	1.1087
1.04	.8508	.1492	.0772	1.1172
1.05	.8531	.1469	.0757	1.1257

(continúa)

TABLA A-4 (*continuación*)

Variable estandarizada	Probabilidades		Esperanzas parciales	
z	$F(z)$	$1 - F(z)$	$L(z)$	$L(-z)$
1.06	.8554	.1446	.0742	1.1342
1.07	.8577	.1423	.0728	1.1428
1.08	.8599	.1401	.0714	1.1514
1.09	.8621	.1379	.0700	1.1600
1.10	.8643	.1357	.0686	1.1686
1.11	.8665	.1335	.0673	1.1773
1.12	.8686	.1314	.0659	1.1859
1.13	.8708	.1292	.0646	1.1946
1.14	.8729	.1271	.0634	1.2034
1.15	.8749	.1251	.0621	1.2121
1.16	.8770	.1230	.0609	1.2209
1.17	.8790	.1210	.0596	1.2296
1.18	.8810	.1190	.0584	1.2384
1.19	.8830	.1170	.0573	1.2473
1.20	.8849	.1151	.0561	1.2561
1.21	.8869	.1131	.0550	1.2650
1.22	.8888	.1112	.0538	1.2738
1.23	.8907	.1093	.0527	1.2827
1.24	.8925	.1075	.0517	1.2917
1.25	.8943	.1057	.0506	1.3006
1.26	.8962	.1038	.0495	1.3095
1.27	.8980	.1020	.0485	1.3185
1.28	.8997	.1003	.0475	1.3275
1.29	.9015	.0985	.0465	1.3365
1.30	.9032	.0968	.0455	1.3455
1.31	.9049	.0951	.0446	1.3446
1.32	.9066	.0934	.0436	1.3636
1.33	.9082	.0918	.0427	1.3727
1.34	.9099	.0901	.0418	1.3818
1.35	.9115	.0885	.0409	1.3909
1.36	.9131	.0869	.0400	1.4000
1.37	.9147	.0853	.0392	1.4092
1.38	.9162	.0838	.0383	1.4183
1.39	.9177	.0823	.0375	1.4275
1.40	.9192	.0808	.0367	1.4367
1.41	.9207	.0793	.0359	1.4459
1.42	.9222	.0778	.0351	1.4551
1.43	.9236	.0764	.0343	1.4643
1.44	.9251	.0749	.0336	1.4736
1.45	.9265	.0735	.0328	1.4828
1.46	.9279	.0721	.0321	1.4921
1.47	.9292	.0708	.0314	1.5014
1.48	.9306	.0694	.0307	1.5107
1.49	.9319	.0681	.0300	1.5200
1.50	.9332	.0668	.0293	1.5293
1.51	.9345	.0655	.0286	1.5386
1.52	.9357	.0643	.0280	1.5480
1.53	.9370	.0630	.0274	1.5574
1.54	.9382	.0618	.0267	1.5667
1.55	.9394	.0606	.0261	1.5761
1.56	.9406	.0594	.0255	1.5855
1.57	.9418	.0582	.0249	1.5949
1.58	.9429	.0571	.0244	1.6044
1.59	.9441	.0559	.0238	1.6138

TABLA A-4 ***(continuación)***

Variable estandarizada	**Probabilidades**		**Esperanzas parciales**	
z	***F*(*z*)**	**1 − *F*(*z*)**	***L*(*z*)**	***L*(−*z*)**
1.60	.9460	.0540	.0232	1.6232
1.61	.9463	.0537	.0227	1.6327
1.62	.9474	.0526	.0222	1.6422
1.63	.9484	.0516	.0216	1.6516
1.64	.9495	.0505	.0211	1.6611
1.65	.9505	.0495	.0206	1.6706
1.66	.9515	.0485	.0201	1.6801
1.67	.9525	.0475	.0197	1.6897
1.68	.9535	.0465	.0192	1.6992
1.69	.9545	.0455	.0187	1.7087
1.70	.9554	.0446	.0183	1.7183
1.71	.9564	.0436	.0178	1.7278
1.72	.9573	.0427	.0174	1.7374
1.73	.9582	.0418	.0170	1.7470
1.74	.9591	.0409	.0166	1.7566
1.75	.9599	.0401	.0162	1.7662
1.76	.9608	.0392	.0158	1.7558
1.77	.9616	.0384	.0154	1.7854
1.78	.9625	.0375	.0150	1.7950
1.79	.9633	.0367	.0146	1.8046
1.80	.9641	.0359	.0143	1.8143
1.81	.9649	.0351	.0139	1.8239
1.82	.9656	.0344	.0136	1.8436
1.83	.9664	.0336	.0132	1.8432
1.84	.9671	.0329	.0129	1.8529
1.85	.9678	.0322	.0126	1.8626
1.86	.9685	.0314	.0123	1.8723
1.87	.9693	.0307	.0119	1.8819
1.88	.9699	.0301	.0116	1.8916
1.89	.9706	.0294	.0113	1.9013
1.90	.9713	.0287	.0111	1.9111
1.91	.9719	.0281	.0108	1.9208
1.92	.9726	.0274	.0105	1.9305
1.93	.9732	.0268	.0102	1.9402
1.94	.9738	.0262	.0100	1.9500
1.95	.9744	.0256	.0097	1.9597
1.96	.9750	.0250	.0094	1.9694
1.97	.9756	.0244	.0092	1.9792
1.98	.9761	.0239	.0090	1.9890
1.99	.9767	.0233	.0087	1.9987
2.00	.9772	.0228	.0085	2.0085
2.01	.9778	.0222	.0083	2.0183
2.02	.9783	.0217	.0080	2.0280
2.03	.9788	.0212	.0078	2.0378
2.04	.9793	.0207	.0076	2.0476
2.05	.9798	.0202	.0074	2.0574
2.06	.9803	.0197	.0072	2.0672
2.07	.9808	.0192	.0072	2.0770
2.08	.9812	.0188	.0068	2.0868
2.09	.9817	.0183	.0066	2.0966
2.10	.9821	.0179	.0065	2.1065
2.11	.9826	.0174	.0063	2.1163
2.12	.9830	.0170	.0061	2.1261
2.13	.9834	.0166	.0060	2.1360
2.14	.9838	.0162	.0058	2.1458

(continúa)

TABLA A-4 (*continuación*)

Variable estandarizada	Probabilidades		Esperanzas parciales	
z	$F(z)$	$1 - F(z)$	$L(z)$	$L(-z)$
2.15	.9842	.0158	.0056	2.1556
2.16	.9846	.0154	.0055	2.1655
2.17	.9850	.0150	.0053	2.1753
2.18	.9854	.0146	.0052	2.1852
2.19	.9857	.0143	.0050	2.1950
2.20	.9861	.0139	.0049	2.2049
2.21	.9864	.0136	.0048	2.2148
2.22	.9868	.0132	.0046	2.2246
2.23	.9871	.0129	.0045	2.2345
2.24	.9875	.0125	.0044	2.2444
2.25	.9878	.0122	.0042	2.2542
2.26	.9881	.0119	.0041	2.2641
2.27	.9884	.0116	.0040	2.2740
2.28	.9887	.0113	.0039	2.2839
2.29	.9890	.0110	.0038	2.2938
2.30	.9893	.0107	.0037	2.3037
2.31	.9896	.0104	.0036	2.3136
2.32	.9898	.0102	.0035	2.3235
2.33	.9901	.0099	.0034	2.3334
2.34	.9904	.0096	.0033	2.3433
2.35	.9906	.0094	.0032	2.3532
2.36	.9909	.0091	.0031	2.3631
2.37	.9911	.0089	.0030	2.3730
2.38	.9913	.0087	.0029	2.3829
2.39	.9916	.0084	.0028	2.3928
2.40	.9918	.0082	.0027	2.4027
2.41	.9920	.0080	.0026	2.4126
2.42	.9922	.0078	.0026	2.4226
2.43	.9925	.0075	.0025	2.4325
2.44	.9927	.0073	.0024	2.4424
2.45	.9929	.0071	.0023	2.4523
2.46	.9931	.0069	.0023	2.4623
2.47	.9932	.0068	.0022	2.4722
2.48	.9934	.0066	.0021	2.4821
2.49	.9936	.0064	.0021	2.4921
2.50	.9938	.0062	.0020	2.5020
2.51	.9940	.0060	.0019	2.5119
2.52	.9941	.0059	.0019	2.5219
2.53	.9943	.0057	.0018	2.5318
2.54	.9945	.0055	.0018	2.5418
2.55	.9946	.0054	.0017	2.5517
2.56	.9948	.0052	.0017	2.5617
2.57	.9949	.0051	.0016	2.5716
2.58	.9951	.0049	.0016	2.5816
2.59	.9952	.0048	.0015	2.5915
2.60	.9953	.0047	.0015	2.6015
2.61	.9955	.0045	.0014	2.6114
2.62	.9956	.0044	.0014	2.6214
2.63	.9957	.0043	.0013	2.6313
2.64	.9959	.0041	.0013	2.6413
2.65	.9960	.0040	.0012	2.6512
2.66	.9961	.0039	.0012	2.6612
2.67	.9962	.0038	.0012	2.6712
2.68	.9963	.0037	.0011	2.6811
2.69	.9964	.0036	.0011	2.6911
2.70	.9965	.0035	.0011	2.7011

TABLA A-4 ***(continuación)***

Variable estandarizada	Probabilidades		Esperanzas parciales	
z	$F(z)$	$1 - F(z)$	$L(z)$	$L(-z)$
2.71	.9966	.0034	.0010	2.7110
2.72	.9967	.0033	.0010	2.7210
2.73	.9968	.0032	.0010	2.7310
2.74	.9969	.0031	.0009	2.7409
2.75	.9970	.0030	.0009	2.7509
2.76	.9971	.0029	.0009	2.7609
2.77	.9972	.0028	.0008	2.7708
2.78	.9973	.0027	.0008	2.7808
2.79	.9974	.0026	.0008	2.7908
2.80	.9974	.0026	.0008	2.8008
2.81	.9975	.0025	.0007	2.8107
2.82	.9976	.0024	.0007	2.8207
2.83	.9977	.0023	.0007	2.8307
2.84	.9977	.0023	.0007	2.8407
2.85	.9978	.0022	.0006	2.8506
2.86	.9979	.0021	.0006	2.8606
2.87	.9979	.0021	.0006	2.8706
2.88	.9980	.0020	.0006	2.8806
2.89	.9981	.0019	.0006	2.8906
2.90	.9981	.0019	.0005	2.9005
2.91	.9982	.0018	.0005	2.9105
2.92	.9982	.0018	.0005	2.9205
2.93	.9983	.0017	.0005	2.9305
2.94	.9984	.0016	.0005	2.9405
2.95	.9984	.0016	.0005	2.9505
2.96	.9985	.0015	.0004	2.9604
2.97	.9985	.0015	.0004	2.9704
2.98	.9986	.0014	.0004	2.9804
2.99	.9986	.0014	.0004	2.9904
3.00	.9986	.0014	.0004	3.0004

Fuente: R.G. Brown, *Decision Rules for Inventory Management* (Hinsdale, IL.; Dryden, 1967), Adaptado de la tabla VI, pp. 95-103.

TABLA A-5 **Factores d_2 para gráficas R**

Cantidad de observaciones en el subgrupo, n	Factor d_2, $d_2 = \frac{\overline{R}}{\sigma}$
2	1.128
3	1.693
4	2.059
5	2.326
6	2.534
7	2.704
8	2.847
9	2.970
10	3.078
11	3.173
12	3.258
13	3.336
14	3.407
15	3.472
16	3.532
17	3.588
18	3.640
19	3.689
20	3.735
21	3.778
22	3.819
23	3.858
24	3.895
25	3.931
30	4.086
35	4.213
40	4.322
45	4.415
50	4.498
55	4.572
60	4.639
65	4.699
70	4.755
75	4.806
80	4.854
85	4.898
90	4.939
95	4.978
100	5.015

Nota: En estos factores se supone muestreo de una población normal.

Fuente: E.L. Grant y R.S. Leavenworth, *Statistical Quality Control*, 6a. edición (Nueva York: McGraw-Hill, 1988).

TABLA A-6 **Factores d_3 y d_4 para gráficas *R***

Cantidad de observaciones en el subgrupo, n	Factores para gráfica *R*	
	Límite inferior de control d_3	Límite superior de control d_4
2	0	3.27
3	0	2.57
4	0	2.28
5	0	2.11
6	0	2.00
7	.08	1.92
8	.14	1.86
9	.18	1.82
10	.22	1.78
11	.26	1.74
12	.28	1.72
13	.31	1.69
14	.33	1.67
15	.35	1.65
16	.36	1.64
17	.38	1.62
18	.39	1.61
19	.40	1.60
20	.41	1.59

Nota: En estos factores se supone muestreo de una población normal. Los límites inferior y superior de control pueden obtenerse de la siguiente manera:

$$\text{LIC} = d_3 R$$
$$\text{LSC} = d_4 R$$

Fuente: E.L. Grant y R.S. Leavenworth, *Statistical Quality Control*, 6a. edición (Nueva York: McGraw-Hill, 1988).

Índice

B

C

F

G

H

K

L

M

N

O

P

S

T

U

V